

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:49 ; Search time 4.53125 Seconds
(without alignments)
106.117 Million cell updates/sec

Title: US-09-829-495-61

Perfect score: 31

Sequence: 1 SYMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	98	2	PL0123
2	31	100.0	101	2	S26460
3	31	100.0	110	2	PH1655
4	31	100.0	114	2	S36280
5	31	100.0	115	2	S38714
6	31	100.0	117	2	S78486
7	31	100.0	117	2	S17079
8	31	100.0	118	2	S00700
9	31	100.0	122	2	A33989
10	31	100.0	122	2	S31675
11	31	100.0	136	2	S31587
12	31	100.0	139	2	I37781
13	31	100.0	147	2	I37780
14	31	100.0	207	2	P95966
15	31	100.0	939	2	AE2275
16	31	100.0	942	2	B84469
17	31	100.0	942	2	T37539
18	31	100.0	1254	1	A32686
19	31	100.0	1844	1	TS1890
20	28	90.3	101	2	H37262
21	28	90.3	103	2	PH0986
22	28	90.3	106	2	S2615
23	28	90.3	108	2	S26316
24	28	90.3	110	2	S26317
25	28	90.3	117	2	G45722
26	28	90.3	118	2	C30560
27	28	90.3	120	2	PD0008
28	28	90.3	126	2	E71185
29	28	90.3	161	2	S49488

30	28	90.3	180	2	T39395	hypothetical oligo
31	28	90.3	190	2	A45601	22k antigen - fluk
32	28	90.3	190	2	A54518	22.6k tegument ant
33	28	90.3	191	2	S41178	gene s32 protein
34	28	90.3	204	2	B82166	conserved hypotet
35	28	90.3	216	2	A72291	hypothetical prote
36	28	90.3	220	2	B49736	collagen alpha 3(I)
37	28	90.3	226	2	T29404	hypothetical prote
38	28	90.3	246	2	I48302	collagen alpha 3(I)
39	28	90.3	253	2	I48304	collagen alpha 5(I)
40	28	90.3	254	2	E83619	probable transport
41	28	90.3	258	2	B61228	collagen alpha 1(I)
42	28	90.3	261	2	A34476	collagen alpha 2(I)
43	28	90.3	309	2	S43573	C05B5.4 protein (c
44	28	90.3	312	2	I48303	collagen alpha 4(I)
45	28	90.3	314	2	AC2690	conserved hypotet

ALIGNMENTS

RESULT 1
PL0123
Ig heavy chain V-III region (TD-Vr) - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
C/Accession: PL0123; S26897
R/Bird, J.; Galli, N.; Link, M.; Srites, D.; Sklar, J.
J. Exp. Med. 168, 229-245, 1988
A/Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin
A/Reference number: PL0116; MUID:88286083; PMID:2840480
A/Accession: PL0123
A/Molecule type: mRNA
A/Residues: 1-98 <RIR>
A/Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
A/Note: The sequence shows the V region (TD-Vr) from a nonproductive DNA rearrangement
R/Tomlinson, I.M.; Walter, G.; Marks, J.D.; Ilewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A/Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A/Reference number: S26885; MUID:93021117; PMID:1404388
A/Accession: S26897
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-98 <ROM>
A/Cross-references: EMBL:Z12354; NID:G32930; PIDN:CAA78224.1; PID:G32931
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: acute lymphoblastic leukemia; heterodimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMV>
F:31-35/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
Db 31 SYMS 35

RESULT 2
S26460
Ig heavy chain V region - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C/Accession: S26460
R/Kavaler, J.
submitted to the EMBL Data Library, April 1991
A/Reference number: S26459
A/Accession: S26460
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-101 <KAV>

A:Cross-references: EMBL:X59106; NID:G51707; PIDN:CAA41832.1; PID:G51708
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 31; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||
Db 13 SYWMS 17

RESULT 3

PH1655
Ig heavy chain V region (clone 2E8) - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C:Accession: PH1655
R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasse, E.H.
J. Exp. Med. 178, 331-336, 1993

A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphylococcal protein A
A:Reference number: PH1642; MUID:93301610; PMID:8315388
A:Accession: PH1655
A:Molecule type: mRNA
A:Residues: 1-110 <HIL>

A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-90/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||
Db 23 SYWMS 27

RESULT 4

S36280
Ig heavy chain V region (clone alpha-FOG1-A3) - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
C:Accession: S36280
R:Gilfillan, A.D.; Malmyvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.
EMBO J. 12, 725-734, 1993

A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679990
A:Accession: S36280
A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 1-114 <GRI>
A:Cross-references: EMBL:Z18822

C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||
Db 31 SYWMS 35

RESULT 5

S38714
Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000

C:Accession: S38714
R:Climis, A.Y.
submitted to the EMBL Data Library, November 1993
A:Reference number: S38713

A:Accession: S38714
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-115 <CIM>

A:Cross-references: EMBL:X76014; NID:G416092; PIDN:CAA53601.1; PID:G1334076
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-99/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||
Db 30 SYWMS 34

RESULT 6

S78486
Ig heavy chain V region (clone FL13-28) - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 21-Jan-2000
C:Accession: S78486; S31115
R:Raaphorst, F.M.
submitted to the EMBL Data Library, October 1991

A:Reference number: S78486
A:Accession: S78486
A:Molecule type: mRNA
A:Residues: 1-117 <RAA>

A:Cross-references: EMBL:X62965
A:Experimental source: clone FL13-28
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurmat
Eur. J. Immunol. 22, 247-251, 1992

A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complex
A:Reference number: S31104; MUID:92111633; PMID:1730252
A:Accession: S31115
A:Molecule type: mRNA

A:Residues: 1-4, 'L', '6-32', 'G', '34-52', 'E', '54-73', 'K', '75-97' <RAW>
A:Cross-references: EMBL:X62965
A:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||
Db 31 SYWMS 35

RESULT 7

S17079
Ig heavy chain V-gene (clone HHG19) - human

C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C:Accession: S60299; S17079
R:Knepers, R.; Fischer, U.; Rajewsky, K.; Gause, A.
Immunol. Lett. 34, 57-62, 1992

A:Title: Immunoglobulin heavy and light chain gene sequences of a human CD5 positive im
A:Reference number: S60295; MUID:9312853; PMID:1282498
A:Accession: S60299
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <K02>

A:Cross-references: EMBL:X62128; NID:G38340; PIDN:CAA44059.1; PID:G38341
A>Note: the authors did not translate the codons for residues 6, 52, 54, 68, 69, 71, 72

C:Genetics: 16/1
A:introns: 16/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|
|
|
|
Db 50 SYWMS 54

RESULT 8

S00700
Ig heavy chain V region - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 23-Jul-1999
C:Accession: S00700
R:Yamawaki-Kataoka, Y.; Honjo, T.
Nucleic Acids Res. 15, 5888, 1987

A:Title: Nucleotide sequences of variable region segments of the immunoglobulin heavy chain
A:Reference number: S00700; WUID:87289054; PMID:3112743
A:Accession: S00700
A:Molecule type: DNA

A:Residues: 1-118 <YAM>
A:Cross-references: EMBL:Y00380; NID:g64810; PID:CAA68452.1; PID:g64811
A>Note: the sequence was determined from the germline gene

C:Genetics: 15/3
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|
|
|
|
Db 50 SYWMS 54

RESULT 9

A33989
Ig heavy chain V-1-D-J region - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 16-Aug-1996
C:Accession: A33989
R:Hsu, E.; Schwager, J.; Alt, F.W.
Proc. Natl. Acad. Sci. U.S.A. 86, 8010-8014, 1989

A:Title: Evolution of immunoglobulin genes: V-H families in the amphibian Xenopus.
A:Reference number: A33989; WUID:90046727; PMID:2510156
A:Accession: A33989
A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA
A:Residues: 1-122 <HSU>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-100/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|
|
|
|
Db 31 SYWMS 35

RESULT 10

S31675

Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S31675
R:Guinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.
submitted to the EMBL Data Library, June 1992

A:Description: Mechanisms that generate human immunoglobulin diversity operate from the
A:Reference number: S31585
A:Accession: S31675
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-122 <GUI>

A:Cross-references: EMBL:Z14175; NID:g31015; PID:CAA78544.1; PID:g31016
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:21-104/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|
|
|
|
Db 37 SYWMS 41

RESULT 11

S31587
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S31587
R:Guinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.

submitted to the EMBL Data Library, June 1992
A:Description: Mechanisms that generate human immunoglobulin diversity operate from the
A:Reference number: S31585
A:Accession: S31587
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-136 <GUI>

A:Cross-references: EMBL:Z14189; NID:g31005; PID:CAA78558.1; PID:g31006
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:31-114/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|
|
|
|
Db 47 SYWMS 51

RESULT 12

I37781
Ig variable region (VDJ) (clone T21-9) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
C:Accession: I37781; S25475
R:Demaison, C.; Chastagner, P.; Theze, J.; Zouali, M.

A:Title: Somatic diversification in the heavy chain variable region genes expressed by h
A:Reference number: A6876; WUID:9411917; PMID:8290556
A:Accession: I37781
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-139 <RES>

A:Cross-references: EMBL:X67908; NID:g33580; PID:CAA40106.1; PID:g33581
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:35-118/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 139;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 |||||
 Db 51 SYWMS 55

RESULT 13

I37780
 Ig variable region (VDJ) (clone T20-11) - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
 C/Accession: I37780; S25474
 R/Demason, C.; Chateigner, P.; Theze, J.; Zouali, M.
 Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
 A/Title: Somatic diversification in the heavy chain variable region genes expressed by h
 A/Reference number: A36876; MUID:94119917; PMID:8290556
 A/Accession: I37780
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-147 <RES>
 A/Cross-references: EMBL:X67943; NID:933578; PIDN:CAA48130.1; PID:933579
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 F/28-11/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 147;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 |||||
 Db 44 SYWMS 48

RESULT 14

F95966
 Probable aldehyde or xanthine dehydrogenase, iron-sulfur subunit protein [imported] - St
 C/Species: Sinorhizobium meliloti
 C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C/Accession: F95966
 R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A/Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
 A/Reference number: A95842; MUID:21396508; PMID:11481431
 A/Accession: F95966
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-207 <KIR>
 A/Cross-references: GB:AL591985; PIDN:CAC49398.1; PID:915140884; GSPDB:GN00167
 A/Experimental source: strain 1021, megaplasmid pSymB
 R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federle, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A/Authors: Kahn, D.; Kahn, M.L.; Kallman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
 Hebutel, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Weller, D.H.; Wong, K.; Yeh, K.
 A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A/Reference number: A96039; MUID:21368234; PMID:11474104
 A/Contents: annotation
 C/Genetics:
 A/Gene: SMD21558
 A/Genome: plasmid

Query Match 100.0%; Score 31; DB 2; Length 207;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 |||||
 Db 178 SYWMS 182

RESULT 15
 AE2275
 hypothetical protein alr3756 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C/Accession: AE2275
 R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyachu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata,
 DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A/Reference number: AB1807; MUID:21595285; PMID:1175840
 A/Accession: AE2275
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-939 <KIR>
 A/Cross-references: GB:BA000019; PIDN:BA075455.1; PID:917132890; GSPDB:GN00179
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: alr3756

Query Match 100.0%; Score 31; DB 2; Length 939;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 |||||
 Db 931 SYWMS 935

Search completed: January 12, 2004, 07:02:57
 Job time : 7.53125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:49 ; Search time 2.73438 Seconds
(without alignments)
85.992 Million cell updates/sec

Title: US-09-829-495-61

Perfect score: 31
Sequence: 1 SYMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt 41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	1254	1 DPOG_YEAST	P15801 saccharomyc
2	28	90.3	168	1 LSPA_VIBVU	Q8d688 vibrio vuln
3	28	90.3	169	1 LSPA_VIBPA	Q87689 vibrio para
4	28	90.3	180	1 ORN_SCHPO	O94626 schizosacch
5	28	90.3	190	1 TEGU_SCHMA	P14602 schistosoma
6	28	90.3	315	1 YK04_CAEEL	P34292 caenorhabdi
7	28	90.3	342	1 TORT_ECO57	P58358 escherichia
8	28	90.3	342	1 TORT_ECOLI	P38683 escherichia
9	28	90.3	359	1 DNL1_BPT7	P00969 bacterioph
10	28	90.3	385	1 YWEM_BACSU	P39596 bacillus su
11	28	90.3	393	1 Y306_WYCGE	P47548 mycoplasma
12	28	90.3	455	1 YAO6_SCHPO	Q10085 schizosacch
13	28	90.3	471	1 CA34_BOVIN	Q28084 bos taurus
14	28	90.3	583	1 PHBC_AZOCA	O66392 a poly-beta
15	28	90.3	745	1 FSP1_RAT	O63517 rattus norv
16	28	90.3	754	1 CAS4_CANPA	Q28247 canis fami
17	28	90.3	1002	1 YA37_ANASP	P58612 anabaena sp
18	28	90.3	1669	1 CA14_HUMAN	P02462 homo sapien
19	28	90.3	1669	1 CA14_MOUSE	P02463 mus musculu
20	28	90.3	1670	1 CA34_HUMAN	Q01955 homo sapien
21	28	90.3	1685	1 CAS4_HUMAN	P29400 homo sapien
22	28	90.3	1691	1 CA64_HUMAN	Q14031 homo sapien
23	28	90.3	1707	1 CA24_MOUSE	P08122 mus musculu
24	28	90.3	1712	1 CA24_HUMAN	P08572 homo sapien
25	28	90.3	1758	1 CA24_CAEEL	P17140 caenorhabdi
26	28	90.3	1763	1 CA24_AOCSU	P27993 ascaris suu
27	28	90.3	2210	1 RRP1_EBOSM	O66802 ebola virus
28	28	90.3	2397	1 MOKB_SCHPO	Q09854 schizosacch
29	28	90.3	3587	1 SRP2_BACSU	Q04747 bacillus su
30	28	90.3	4540	1 DYHC_PAPTE	Q27171 paramectum
31	27	87.1	116	1 HV36_MOUSE	P01806 mus musculu
32	27	87.1	117	1 HV04_MOUSE	P01748 mus musculu
33	27	87.1	117	1 HV05_MOUSE	P01749 mus musculu

ALIGNMENTS

```

34 27 87.1 117 1 HV06_MOUSE P01750 mus musculu
35 27 87.1 117 1 HV09_MOUSE P01753 mus musculu
36 27 87.1 117 1 HV10_MOUSE P01811 mus musculu
37 27 87.1 117 1 HV41_MOUSE P01812 mus musculu
38 27 87.1 117 1 HV42_MOUSE P06328 mus musculu
39 27 87.1 117 1 HV49_MOUSE P01809 mus musculu
40 27 87.1 118 1 HV39_MOUSE P01807 mus musculu
41 27 87.1 119 1 HV37_MOUSE P01808 mus musculu
42 27 87.1 119 1 HV38_MOUSE P01810 mus musculu
43 27 87.1 119 1 HV40_MOUSE P06329 mus musculu
44 27 87.1 120 1 HV50_MOUSE P01751 mus musculu
45 27 87.1 139 1 HV07_MOUSE

RESULT 1
DPOG_YEAST STANDARD; PRT; 1254 AA.
ID P15801; Q08785;
AC 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase gamma (EC 2.7.7.7) (Mitochondrial DNA polymerase
DE catalytic subunit).
GN MIPI OR YOR330C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90062193; PubMed=2684980;
RA Foury F.;
RT "Cloning and sequencing of the nuclear gene MIPI encoding the
RT catalytic subunit of the yeast mitochondrial DNA polymerase.";
RL J. Biol. Chem. 264:20552-20560(1989).
RN [2]
RP REVISIONS.
RA Foury F.;
RT Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97051586; PubMed=8896263;
RA Parle-Medemont A.G., Hand N.J., Goulding S.G., Wolfe K.H.;
RT "Sequence of 29 kb around the PDR10 locus on the right arm of
RT Saccharomyces cerevisiae chromosome XV: similarity to part of
RT chromosome I.";
RL Yeast 12:999-1004(1996).
RN [4]
RP SEQUENCE OF 987-1254 FROM N.A.
RC STRAIN=GRF88;
RA Song J.M., Cheung E., Rabinowitz J.C.;
RT Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION: INVOLVED IN THE REPLICATION OF MITOCHONDRIAL DNA.
CC - FUNCTIONAL ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC - COFACTOR: MAGNESTIUM.
CC - SUBCELLULAR LOCATION: Mitochondrial.
CC - MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
CC DIFFERENT REACTIONS OF DNA SYNTHESIS.
CC - SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR EMBL; J05117; AAA17543.1; -
DR EMBL; Z49821; CAA89977.1; -
DR EMBL; Z75238; CAA99652.1; ALT_INIT.
DR EMBL; U42227; AAA85442.1; -
DR PIR; S62062; A32686.
DR SGD; S0005857; MIP1.
DR GO; GO:0005739; C:mitochondrion; IDA.
DR GO; GO:0006261; P:DNA dependent DNA replication; IDA.
DR InterPro; IPR001098; DNA_pol.
DR InterPro; IPR002297; DNA_polG.
DR Pfam; PF00476; DNA_POL_A; 1.
DR PRINTS; PR00867; DNAPOLG.
DR SMART; SM00482; POLAC; 1.
DR PROSITE; PS00447; DNA POLYMERASE A; 1.
DR TransErase; DNA-directed DNA polymerase; DNA replication;
KM DNA-binding; Mitochondrion; Magnesium.
FT DOMAIN 150 189 CONSENSUS FOR POLYMERASES MOTIF II.
FT DOMAIN 198 198 CONSENSUS FOR POLYMERASES MOTIF I.
FT DOMAIN 225 255 CONSENSUS FOR POLYMERASES MOTIF VI.
FT CONFLICT 8 8 S -> F (IN REF. 1).
FT CONFLICT 35 35 T -> A (IN REF. 1).
FT CONFLICT 222 222 I -> V (IN REF. 1).
FT CONFLICT 357 357 E -> K (IN REF. 1).
FT CONFLICT 540 541 TH -> MN (IN REF. 1).
FT CONFLICT 616 616 S -> N (IN REF. 1).
FT CONFLICT 661 661 A -> T (IN REF. 1).
FT CONFLICT 978 978 S -> P (IN REF. 1).
FT CONFLICT 986 986 N -> S (IN REF. 1).
SQ SEQUENCE 1254 AA; 143501 MW; 80BE686BE8FA9F2EC CRC64;

Query Match Best Local Similarity 100.0%; Score 31; DB 1; Length 1254;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
Db 599 SYMMS 603

RESULT 2
LSPA_VIBVU STANDARD; PRT; 168 AA.
ID LSPA_VIBVU
AC Q87889;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lipoprotein signal peptidase (EC 3.4.23.36) (Prolipoprotein signal
peptidase) (Signal peptidase II) (SPase II).
GN LSPA OR V10506.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
Choi H.E.;
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RL FUNCTION: This protein specifically catalyzes the removal of
signal peptides from prolipoproteins (By similarity).
CC -1- CATALYTIC ACTIVITY: Cleavage of N-terminal leader sequences from
membrane prolipoproteins. Hydrolyses Xaa-Xbb-Xcc-|-Cys, in which
Xaa is hydrophobic (preferably Leu), Xbb is often Ser or Ala, Xcc
is often Gly or Ala, and the Cys is alkylated on sulfur with a
diacylglyceryl group.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Probable).
CC -1- SIMILARITY: Belongs to peptidase family A8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF016798; AAO09025.1; -
CC HAMAP; MF_00161; -; 1.
CC InterPro; IPR001872; SigPase_A8.
CC Pfam; PF01252; Peptidase A8; 1.
CC PRINTS; PR00781; LIPOSIGPASE.
CC PRODOM; PD004304; SigPase_A8; 1.
CC TIGRfam; TIGR00077; lspa_1.
CC DR PROSITE; PS00855; SPASE_II; 1.
CC Hydrolyase; Aspartyl protease; Transmembrane; Inner membrane;
KM Complete proteome.
FT TRANSMEM 12 34 Potential.
FT TRANSMEM 70 92 Potential.
FT TRANSMEM 104 126 Potential.
FT TRANSMEM 136 158 Potential.
FT ACT SITE 119 119 BY SIMILARITY.
FT ACT SITE 146 146 BY SIMILARITY.
SQ SEQUENCE 168 AA; 19200 MW; 82529709B2D7BA3 CRC64;

Query Match Best Local Similarity 90.3%; Score 28; DB 1; Length 168;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
Db 89 TYMMS 93

RESULT 3
LSPA_VIBPA STANDARD; PRT; 169 AA.
ID LSPA_VIBPA
AC Q87889;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lipoprotein signal peptidase (EC 3.4.23.36) (Prolipoprotein signal
peptidase) (Signal peptidase II) (SPase II).
GN LSPA OR VP0535.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; Pubmed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
Tajima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
Yasunaga T., Honda T., Shinagawa H., Hattori W., Iida T.,
RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
distinct from that of *V. cholerae*.";
RL Lancet 361:743-749(2003).
CC -1- FUNCTION: This protein specifically catalyzes the removal of
signal peptides from prolipoproteins (By similarity).
CC -1- CATALYTIC ACTIVITY: Cleavage of N-terminal leader sequences from
membrane prolipoproteins. Hydrolyses Xaa-Xbb-Xcc-|-Cys, in which
Xaa is hydrophobic (preferably Leu), Xbb is often Ser or Ala, Xcc
is often Gly or Ala, and the Cys is alkylated on sulfur with a
diacylglyceryl group.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Probable).
CC -1- SIMILARITY: Belongs to peptidase family A8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AP005074; BAC58798.1; -

CC DR HAMAP; MF 00161; -; 1.

CC DR PROSITE; P50085; SPASE_1; 1.

CC KW Hydrolyase; Aspartyl protease; Transmembrane; Inner membrane;

CC Complete proteome.

CC FT TRANSMEM 15 34 Potential.

CC FT TRANSMEM 47 66 Potential.

CC FT TRANSMEM 70 92 Potential.

CC FT TRANSMEM 104 126 Potential.

CC FT TRANSMEM 136 158 Potential.

CC FT ACT SITE 119 119 BY SIMILARITY.

CC FT ACT SITE 146 146 BY SIMILARITY.

CC FT ACT SITE 146 146 BY SIMILARITY.

CC SQ SEQUENCE 169 AA; 19265 MW; F9FA101EC910A74F CRC64;

Query Match 90.3%; Score 28; DB 1; Length 169;

Best local Similarity 80.0%; Pred. No. 1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5

DB 89 TYWMS 93

RESULT 4

ORF SCHPO STANDARD; PRT; 180 AA.

ID ORF SCHPO

AC 094636;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable oligoribonuclease (EC 3.1.-.-).

GN SPBC1347.07.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21648401; PubMed=11859360;

RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbinowitsch E.,

RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Stevens S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Wolckstein G., Xert R., Roben U., Grynolprez B.,

RA Welfens J., Vanstreels E., Rieger M., Schaefer M., Hubert-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hubert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Medler H., Manbuit R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,

RA Galibert F., Aves S., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovskii G.V., Ussery D., Barrell B.G., Nurse P.,

RT "The genome sequence of Schizosaccharomycetes pombe."

RL Nature 415:871-880(2002).

CC -1- FUNCTION: 3'-TO-5' EXORIBONUCLEASE SPECIFIC FOR SMALL

CC OLIGORIBONUCLEOTIDES (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE OLIGORIBONUCLEASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AL035548; CAB37438.1; -

CC DR PIR; T39395; T39395.1; -

CC DR GeneDB; SP006; SPBC1347.07; -

CC DR InterPro; IPR006055; Exonuclease.

CC DR Pfam; PF00929; Exonuclease; 1.

CC DR SMART; SM00479; EXOIII; 1.

CC KW Hypothetical protein; Hydrolyase; Exonuclease; Nuclease.

CC ACT SITE 130 130 POTENTIAL.

CC SQ SEQUENCE 180 AA; 20785 MW; 7AE466905131B603 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 180;

Best local Similarity 80.0%; Pred. No. 1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5

DB 176 TYWMS 180

RESULT 5

TEGU SCHMA STANDARD; PRT; 190 AA.

ID TEGU SCHMA

AC P14202;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Tegument antigen (I(HA) (Antigen SMA22.6) (A12).

OS Schistosoma mansoni (Blood fluke).

OC Schistosoma mansoni (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoidae;

OC Schistosomatidae; Schistosomatidae; Schistosoma.

OX NCBI_TaxID=6183;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=Puerto Rican;

RX MEDLINE=87014570; PubMed=2429181;

RA Stein L.D., David J.R.,

RT "Cloning of a developmentally regulated tegument antigen of

RT Schistosoma mansoni."

RL Mol. Biochem. Parasitol. 20:253-264(1986).

RN (2)

RP SEQUENCE FROM N.A.

RX MEDLINE=91304514; PubMed=1852171;

RA Ueffe S.A., Hagan P., Allen R., Correa-Oliveira R., Smithers S.R.,

RA Simpson A.J.G.;

RT "Molecular cloning and characterisation of the 22-kilodalton adult

RT Schistosoma mansoni antigen recognised by antibodies from mice

RT protectively vaccinated with isolated tegumental surface membranes."

RL Mol. Biochem. Parasitol. 46:159-168(1991).

CC -1- TISSUE SPECIFICITY: ADULT TEGUMENT.

CC -1- DEVELOPMENTAL STAGE: THIS ANTIGEN OCCURS IN ADULTS & SPOROZYSTS

CC BUT NOT IN CERCARIAE, EGGS OR NEWLY TRANSFORMED SCHISTOSOMULA.

CC IT IS A DEVELOPMENTALLY REGULATED PROTEIN.

CC -1- SIMILARITY: TO S.MANSONI ANTIGEN SM21.7.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M29837; AAA29856.1; -

CC DR EMBL; M37003; AAA29922.1; -

DR PIR; A54518; A54518.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 2.
 DR SMART; SM00054; Eph; 2.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
 KW Antigen.
 SQ SEQUENCE 190 AA; 22578 MW; F265EFA29AA7ACDB CRC64;

Query Match 90.3%; Score 28; DB 1; Length 190;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMMS 5
 Db 160 SYMMS 164

RESULT 6
 YKO4_CAEEL STANDARD; PRT; 315 AA.

AC P34252;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Hypothetical 35.9 kDa protein C05B5.4 in chromosome III.
 GN C05B5.4.
 OS Caenorhabditis elegans.
 OC Bkaryotes; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN 1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Mortimore B.J.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z32679; CAAB3591.1; -
 DR PIR; B88571; B88571.
 DR WormPep; C05B5.4; CE00880.
 KW Hypothetical protein.
 SQ SEQUENCE 315 AA; 35918 MW; D665F5FAB040BCF CRC64;

Query Match 90.3%; Score 28; DB 1; Length 315;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMMS 5
 Db 281 SYMMS 285

RESULT 7
 TORT_ECO57 STANDARD; PRT; 342 AA.

AC P58356;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Periplasmic protein tort precursor.
 GN TORT OR Z1411 OR ECG1149.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN 1)

RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobleck E.J., Davis N.W., Lim A., Dinalanta E.T., Potamowski K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533 (2001).

RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22(2001).

CC -1- FUNCTION: Upon binding a putative inducer it probably interacts
 CC with tors and allows it to play a role in the induction of the
 CC torCAD operon for trimethylamine N-oxide reductase (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
 CC RECEPTOR FAMILY 2.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; AB005293; AAG55541.1; -
 DR PIR; AP002554; BAB34572.1; -
 DR PIR; A85635; A85635.
 DR PIR; E90772; E90772.
 DR InterPro; IPR001761; PeriplBP/Lact.
 DR Pfam; PF00532; Peripla_BP_1like; 1.
 DE Transport; Periplasmic; Signal; Complete proteome.
 FT SIGNAL 1 BY SIMILARITY.
 FT CHAIN 19 342 PERIPLASMIC PROTEIN TORT.
 FT SEQUENCE 342 AA; 37950 MW; 8B44A1B064D8A54D CRC64;

Query Match 90.3%; Score 28; DB 1; Length 342;
 Best Local Similarity 80.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMMS 5
 Db 57 SYMMS 61

RESULT 8
 TORT_ECOLI STANDARD; PRT; 342 AA.

AC P38683; P75888;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Periplasmic protein tort precursor.
 GN TORT OR B0994.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN 1)
 RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MC4100;
 RA MEDLINE=94364937; PubMed=8083154;
 RA Simon G., Mejean V., Jourlin C., Chippaux M., Pascal M.-C.;
 RT "The torR gene of *Escherichia coli* encodes a response regulator
 RT protein involved in the expression of the trimethylamine N-oxide
 RT reductase genes."
 RT J. Bacteriol. 176:5601-5606(1994).
 RN (2)
 RP REVISIONS TO 29-30, SEQUENCE OF 19-24, AND CHARACTERIZATION.
 RC STRAIN=K12;
 RA MEDLINE=96165284; PubMed=8576063;
 RA Jourlin C., Simon G., Pommier J., Chippaux M., Mejean V.;
 RT "The periplasmic TorT protein is required for trimethylamine N-oxide
 RT reductase gene induction in *Escherichia coli*."
 RT J. Bacteriol. 178:1219-1223(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RA MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of *Escherichia coli* K-12."
 RT Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayaishi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horinouchi T.;
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map."
 RT DNA Res. 3:137-155(1996).
 CC -1- FUNCTION: Upon binding a putative inducer it probably interacts
 CC with torS and allows it to play a role in the induction of the
 CC torCAB operon for trimethylamine N-oxide reductase.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
 CC RECEPTOR FAMILY 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X94231; CAA63921.1; -;
 DR EMBL; AE000201; AAC74079.1; -;
 DR EMBL; D90736; BAA36136.1; -;
 DR EMBL; D90737; BAA35761.1; -;
 DR PIR; H64840; H64840.
 DR EcdGene; E612616; torT.
 DR InterPro; IPR001761; Periplasmic/Lact.
 DR Pfam; PF00532; Peripla_BP_like; 1.
 KW Transport; Periplasmic; Signal; Complete proteome.
 FT CHAIN 1 18
 FT SIGNAL 1 18
 FT CHAIN 19 342 PERIPLASMIC PROTEIN TOR T.
 FT CONFLICT 59 59 W -> S (IN REF. 1).
 FT CONFLICT 230 230 A -> R (IN REF. 1).
 SQ SEQUENCE 342 AA; 37864 MW; 82FC7AEDF3B2DB9 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 342;
 Best Local Similarity 80.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5

DB 57 SYWMS 61
 ID DNL1_BPT7 STANDARD; PRT; 359 AA.
 AC P00969;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA ligase (EC 6.5.1.1) (Polydeoxyribonucleotide synthase [ATP]).
 GN 1.3.
 OS Bacteriophage T7.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC T7-like viruses.
 OC NCBI_TaxID=10760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=83241725; PubMed=6864790;
 RA Dunn J.J., Studier F.W.;
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
 RT locations of T7 genetic elements."
 RT J. Mol. Biol. 166:477-535(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=82078034; PubMed=7310871;
 RA Dunn J.J., Studier F.W.;
 RT "Nucleotide sequence from the genetic left end of bacteriophage T7
 RT DNA to the beginning of gene 4."
 RT J. Mol. Biol. 148:303-330(1981).
 RN [3]
 RP SEQUENCE OF 278-359 FROM N.A.
 RA MEDLINE=81053683; PubMed=7001354;
 RA Dunn J.J., Studier F.W.;
 RT "The transcription termination site at the end of the early region of
 RT bacteriophage T7 DNA."
 RT Nucleic Acids Res. 8:2119-2132(1980).
 RN [4]
 RP SEQUENCE OF 1-5 FROM N.A.
 RA MEDLINE=81054683; PubMed=6254001;
 RA Saito H., Tabors S., Tamanoi F., Richardson C.C.;
 RT "Nucleotide sequence of the primary origin of bacteriophage T7 DNA
 RT replication: relationship to adjacent genes and regulatory
 RT elements."
 RT Proc. Natl. Acad. Sci. U.S.A. 77:3917-3921(1980).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RA MEDLINE=96222302; PubMed=8653795;
 RA Subramanya H.S., Doherty A.J., Ashford S.R., Wigley D.B.;
 RT "Crystal structure of an ATP-dependent DNA ligase from bacteriophage
 RT T7."
 RT Cell 85:607-615(1996).
 CC -1- FUNCTION: DNA LIGASE, WHICH IS EXPRESSED IN THE EARLY STAGE OF
 CC LYtic DEVELOPMENT, HAS BEEN IMPLICATED IN T7 DNA SYNTHESIS AND
 CC GENETIC RECOMBINATION. IT MAY ALSO PLAY A ROLE IN T7 DNA REPAIR.
 CC -1- CATALYTIC ACTIVITY: ATP + {deoxyribonucleotide} (N) +
 CC {deoxyribonucleotide} (M) = AMP + diphosphate +
 CC {deoxyribonucleotide} (N+M).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT DNA LIGASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V01124; CAA24322.1; -;
 DR EMBL; V01126; CAA24336.1; -;
 DR EMBL; V01127; CAA24336.1; -;
 DR EMBL; V01146; CAA24393.1; -;

DR PIR; E94615; LOBP37.
 DR PDB; 1A01; 25-MAR-98.
 DR InterPro; IPR000977; DNA_ligase.
 DR Pfam; PF01068; DNA_ligase; 1.
 DR PROSITE; PS00697; DNA_LIGASE_A1; 1.
 DR PROSITE; PS00333; DNA_LIGASE_A2; 1.
 DR PROSITE; PS00160; DNA_LIGASE_A3; 1.
 KM DNA repair; DNA replication; DNA recombination; ATP-binding; ligase;
 3D-structure.
 KM BINDING 34 34 AMP (BY SIMILARITY).
 FT TURN 3 4
 FT STRAND 9 13
 FT HELIX 16 26
 FT STRAND 29 33
 FT STRAND 38 45
 FT TURN 46 47
 FT STRAND 49 53
 FT TURN 55 56
 FT STRAND 59 59
 FT HELIX 61 66
 FT STRAND 68 69
 FT HELIX 71 78
 FT TURN 80 81
 FT TURN 85 86
 FT STRAND 87 96
 FT TURN 101 103
 FT HELIX 104 109
 FT STRAND 117 117
 FT STRAND 139 139
 FT TURN 141 143
 FT STRAND 144 153
 FT HELIX 154 159
 FT STRAND 163 166
 FT HELIX 167 184
 FT TURN 186 187
 FT STRAND 189 192
 FT STRAND 195 198
 FT HELIX 201 212
 FT TURN 213 215
 FT STRAND 219 222
 FT TURN 224 225
 FT STRAND 227 228
 FT STRAND 231 239
 FT STRAND 243 253
 FT TURN 257 259
 FT STRAND 267 271
 FT TURN 273 274
 FT STRAND 277 281
 FT HELIX 286 299
 FT TURN 300 302
 FT TURN 321 324
 FT STRAND 326 331
 FT STRAND 333 334
 FT TURN 336 337
 FT STRAND 340 341
 FT STRAND 344 348
 SO SEQUENCE 359 AA; 41133 MW; 43DA453B71BDF8DC CRC64;

Query Match 90.3%; Score 28; DB 1; Length 359;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
 DB 50 SYMLS 54

RESULT 10
 YWBM_BACSU STANDARD; PRT; 385 AA.
 AC P39596;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ywbm.
 GN YWBM OR IPA-28D.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillu.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95020537; PubMed=7934828;
 RA Glaser P., Kunst F., Arnaut M., Coudart M.P., Gonzales W.,
 RA Hulio M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
 RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
 RA Rapoport G., Danchin A.;
 RT "Bacillus subtilis genome project: cloning and sequencing of the 97
 kb region from 325 degrees to 333 degrees.";
 RL Mol. Microbiol. 10:371-384 (1993).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertorello M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani U.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
 RA Giuseppe G., Guy B.J., Hage K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones U.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetle D., Potwolik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler K., Yasunoto K., Yata K.,
 RA Winters P., Wipet A., Yamamoto H., Yamane K., Weitzenger T.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256 (1997).
 CC -I- SIMILARITY: TO E.COLI YCDO.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; X73124; CA51584.1; -;
 DR EMBL; Z99123; CAB15853.1; -;
 DR PIR; S39683; S39683.
 DR SUDLIS; B610574; ywbm.
 DR Pfam; PF04302; DUF451; 1.
 KM Hypothetical protein; Complete proteome.
 SO SEQUENCE 385 AA; 42796 MW; 1F9DC6BA385A3C7 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 385;
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 DB 78 SYWLS 82

RESULT 11
 Y306 MYCGB STANDARD; PRT; 393 AA.
 AC P47548;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MG306.
 GN MG306.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxId=2097;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Incier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC CC
 CC EMBL; U39711; AAC71528.1; -.
 DR PIR; H64233; H64233.
 DR TIGR; MG306; -.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 56 76 POTENTIAL.
 FT TRANSMEM 86 106 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 176 196 POTENTIAL.
 FT TRANSMEM 253 273 POTENTIAL.
 FT TRANSMEM 289 309 POTENTIAL.
 FT TRANSMEM 349 369 POTENTIAL.
 SQ SEQUENCE 393 AA; 45750 MW; D740FDB979EC364A CRC64;

Query Match 90.3%; Score 28; DB 1; Length 393;
 Best Local Similarity 80.0%; Pred. No. 2.4e+2;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 DB 173 SYWLS 177

RESULT 12
 YAO6 SCHPO STANDARD; PRT; 455 AA.
 AC Q10085;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C1D3.06 in chromosome I.
 GN SPAC1D3.06.
 OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxId=4896;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins W., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsle K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langier I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forstner S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC CC -1- SIMILARITY: BELONGS TO THE MULTI ANTIMICROBIAL EXTRUSION (MATE)
 CC FAMILY.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC CC
 CC EMBL; Z68166; CAA92307.1; -.
 DR PIR; T37517; T37517.
 DR GeneDB SPombe; SPAC1D3.06; -.
 DR InterPro; IPR002528; Mate.
 DR Pfam; PF01554; Mate; 2.
 DR TIGRfam; TIGR00797; mate; 1.
 KW Hypothetical protein; Transmembrane; Transport.
 FT TRANSMEM 4 24 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 87 107 POTENTIAL.
 FT TRANSMEM 159 179 POTENTIAL.
 FT TRANSMEM 184 204 POTENTIAL.
 FT TRANSMEM 236 256 POTENTIAL.
 FT TRANSMEM 280 300 POTENTIAL.
 FT TRANSMEM 309 329 POTENTIAL.
 FT TRANSMEM 343 363 POTENTIAL.
 FT TRANSMEM 381 401 POTENTIAL.
 FT TRANSMEM 405 425 POTENTIAL.
 SQ SEQUENCE 455 AA; 49086 MW; FFA1087BC405327 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 455;
 Best Local Similarity 80.0%; Pred. No. 2.8e+2;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 DB 196 SYWLS 200

RESULT 13
CA34_BOVIN STANDARD; PRT; 471 AA.
AC Q28084;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Collagen alpha 3 (IV) chain (Fragment).
GN COL4A3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lens;
RX MEDLINE=91093146; PubMed=1985905;
RA Morrison K.E., Germino G.G., Reeder S.T.;
RT "Use of the polymerase chain reaction to clone and sequence a cDNA
RT encoding the bovine alpha 3 chain of type IV collagen.";
RL J. Biol. Chem. 266:34-39(1991).
RN [2]
RP SEQUENCE OF 227-258.
RC TISSUE=Kidney;
RX MEDLINE=9020779; PubMed=2318822;
RA Gunwar S., Saus J., Noelken M.E., Hudson B.G.;
RT "Glomerular basement membrane. Identification of a fourth chain,
RT alpha 4, of type IV collagen.";
RL J. Biol. Chem. 265:5466-5469(1990).
RN [3]
RP SEQUENCE OF 227-254.
RX MEDLINE=88330844; PubMed=2417661;
RA Saus J., Wieslander J., Langeveld J.P.M., Quinones S., Hudson B.G.;
RT "Identification of the Goodpasture antigen as the alpha 3 (IV) chain
RT of collagen IV.";
RL J. Biol. Chem. 263:13374-13380(1988).
RN [4]
RP SEQUENCE OF 227-244.
RX MEDLINE=87222419; PubMed=2438283;
RA Butkewski R.J., Langeveld J.P.M., Wieslander J., Hamilton J.,
RA Hudson B.G.;
RT "Localization of the Goodpasture epitope to a novel chain of basement
RT membrane collagen.";
RL J. Biol. Chem. 262:7874-7877(1987).
CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
CC GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
CC MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/
CC NIDDOGEN.
CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC -1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC TRIPLE-HELICAL 7S DOMAIN.
CC -1- PPM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- PPM: Type IV collagens contain numerous cysteine residues which
CC are involved in inter- and intramolecular disulfide bonding. 12 of
CC these, located in the NC1 domain, are conserved in all known type
CC IV collagens.
CC -1- SIMILARITY: TO OTHER TYPE IV COLLAGENS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M63139; AAA62708.1; -.
CC DR PIR; A39024; A39024.
CC DR InterPro; IPR000087; Collagen.
CC DR InterPro; IPR001442; ProcollagenC4.
CC DR Pfam; PF01413; C4; 2.
CC DR Pfam; PF01391; Collagen; 4.
CC DR Prodom; PD003923; Procollagnc4; 1.
CC DR SMART; SM00111; C4; 2.
CC KM Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
CC KW Glycoprotein; Basement membrane; Collagen; Cell adhesion.
CC FT NON_TER 1 238 TRIPLE-HELICAL REGION.
CC FT DOMAIN 239 471 NONHEMICAL REGION (NC1).
CC FT SITE 106 108 CELL ATTACHMENT SITE (POTENTIAL).
CC FT MOD_RES 232 232 HYDROXYLATION.
CC FT DISULFID 261 352 OR 349 (BY SIMILARITY).
CC FT DISULFID 294 349 OR 352 (BY SIMILARITY).
CC FT DISULFID 371 466 BY SIMILARITY.
CC FT DISULFID 405 463 OR 463 (BY SIMILARITY).
CC FT DISULFID 417 423 BY SIMILARITY.
CC FT CONFLICT 253 253 S -> Y (IN REF. 3).
CC SQ SEQUENCE 471 AA; 47585 MW; C03B6F14E7008DE CRC64;
CC
CC Query Match 90.3%; Score 28; DB 1; Length 471;
CC Best Local Similarity 80.0%; Pred. No. 2.9e+02;
CC Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 SYMMS 5
CC Db 321 SYMMS 325
CC
CC RESULT 14
CC PHBC_AZOCA STANDARD; PRT; 583 AA.
AC 066352;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Poly-beta-hydroxybutyrate polymerase (EC 2.3.1.-) (Poly(3-
DE hydroxybutyrate) polymerase) (PHB polymerase) (PHA synthase)
DE (Poly(hydroxyalkanoate) polymerase) (PHA synthase).
GN PHBC.
OS Azorhizobium caulinodans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Hyphomicrobiaceae; Azorhizobium.
OX NCBI_TaxId=7;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORS571;
RX MEDLINE=98422458; PubMed=9748438;
RA Mandon K., Michel-Reyelle N., Encarnacion S., Kaminski P.A.,
RA Leija A., Cavallo M.A., Elmerich C., More J.;
RT "Poly-beta-hydroxybutyrate turnover in Azorhizobium caulinodans is
RT required for growth and affects nifH expression.";
RL J. Bacteriol. 180:5070-5076(1998).
CC -1- FUNCTION: POLYMERIZES D(-)-3-HYDROXYBUTYRYL-COA TO CREATE PHB
CC WHICH CONSISTS OF THOUSANDS OF HYDROXYBUTYRATE RESIDUES LINKED
CC END TO END. PHB SERVES AS AN INTRACELLULAR ENERGY RESERVE MATERIAL
CC WHEN CELLS GROW UNDER CONDITIONS OF NUTRIENT LIMITATION.
CC -1- PATHWAY: Poly-beta-hydroxybutyrate biosynthesis; third step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PHA/PHB SYNTHASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

Search completed: January 12, 2004, 06:59:35
 Job time : 5.73438 secs

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AJ006237; CAA06928.1; -
 KW PHB biosynthesis; Transferase; Acyltransferase.
 FT ACT SITE 320 POTENTIAL.
 SQ SEQUENCE 583 AA; 64739 MW; C205763D6DD1A18 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 583;
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 :|||
 Db 50 NYWMS 54

RESULT 15

FSPL_RAT STANDARD; PRT; 745 AA.
 ID FSPL_RAT
 AC 063517;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Leucine-rich primary response protein 1 (Follicle-stimulating hormone
 DE primary response protein).
 GN FSHRH1 OR LRPL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Miscar; TISSUE=Testis;
 RX MEDLINE=95278605; PubMed=7758824;
 RA Slegtenhorst-Begedeman K.E., Post M., Baarends W.M., Themmen A.P.M.,
 RA Grootegeed J.A.;
 RT "Regulation of gene expression in Sertoli cells by
 follicle-stimulating hormone (FSH): cloning and characterization of
 LRPL, a primary response gene encoding a leucine-rich protein.";
 RL Mol. Cell. Endocrinol. 108:115-124(1995).
 CC -1- FUNCTION: INVOLVED IN THE RESPONSE OF GONADAL TISSUES TO FOLLICLE-
 CC STIMULATING HORMONE.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS, OVARY AND SPLEEN.
 CC A MUCH LOWER MENA LEVEL IS FOUND IN BRAIN AND LUNG, AND NO
 CC EXPRESSION IS DETECTED IN LIVER, KIDNEY, HEART, MUSCLE, PITUITARY
 CC GLAND, PROSTATE, EPIDIDYMIS AND SEMINAL VESICLE.
 CC -1- INDUCTION: BY FOLLICLE-STIMULATING HORMONE (FSH).
 CC -1- SIMILARITY: SOME, TO S.POMBE MTS6.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X90355; CAA62018.1; -
 CC PIR; I57665; I57665.
 DR SEQUENCE 745 AA; 85650 MW; 3C9D1AD066D4FD85 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 745;
 Best Local Similarity 80.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 :|||
 Db 383 SYWMS 387

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:04 ; Search time 10.1562 seconds
(without alignments)
127.041 Million cell updates/sec

Title: US-09-829-495-61
Perfect score: 31
Sequence: 1 SYWMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP archaea:*
- 2: SP bacteria:*
- 3: SP fungi:*
- 4: SP human:*
- 5: SP_invertebrate:*
- 6: SP_mammal:*
- 7: SP_mhc:*
- 8: SP organelle:*
- 9: SP phage:*
- 10: SP plant:*
- 11: SP rodent:*
- 12: SP_virus:*
- 13: SP vertebrate:*
- 14: SP_unclassified:*
- 15: SP_virus:*
- 16: SP_bacteriap:*
- 17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	31	100.0	95	4	Q9ULB6	Q9ulb6 homo sapien
2	31	100.0	207	16	Q92UT9	Q92ut9 rhizobium m
3	31	100.0	395	16	Q8NRH2	Q8nrh2 corynebacte
4	31	100.0	399	16	Q8FOF6	Q8fof6 corynebacte
5	31	100.0	939	16	Q8YOR3	Q8yor3 anabaena sp
6	31	100.0	942	3	O13690	O13690 schizosacch
7	31	100.0	942	10	Q9SLI8	Q9sl18 arabidopsis
8	28	90.3	57	8	Q955K1	Q955k1 mytilus tro
9	28	90.3	89	8	Q8LXR7	Q8lxr7 mytilus tro
10	28	90.3	105	8	Q8LXS2	Q8lxs2 mytilus edu
11	28	90.3	107	8	Q8LXS3	Q8lxs3 mytilus edu
12	28	90.3	108	8	Q8LXS0	Q8lxs0 mytilus edu
13	28	90.3	108	8	Q8LXS1	Q8lxs1 mytilus edu
14	28	90.3	108	8	Q8LXA4	Q8lxa4 mytilus edu
15	28	90.3	108	8	Q8LXR9	Q8lxr9 mytilus edu
16	28	90.3	108	8	Q8LVR8	Q8lvr8 mytilus edu

ALIGNMENTS

RESULT 1	ID	Q9ULB6	PRELIMINARY;	PRT;	95 AA.
AC	Q9ULB6	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-MAY-2003 (TREMBLrel. 23, Last annotation update)				
DE	Immunoglobulin heavy chain (Fragment).				
GN	VH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Tange Y., Kayano H.;				
RT	"Human VH gene sequence."				
RL	Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; AB035268; BAA87067.1; ..				
DR	HSSP; P01772; 2FB4.				
DR	InterPro; IPR007110; IG_1-like.				
DR	InterPro; IPR003006; IG_MHC.				
DR	InterPro; IPR003596; IG_V.				
DR	Pfam; PF00447; IG_1.				
DR	SMART; SM00406; IGV; 1.				
DR	PROSITE; PS50835; IG_LIKE; 1.				
FT	NON_TER	1			
FT	NON_TER	1			
SQ	SEQUENCE	95 AA; 10527 MM; 90A86D16D22574A CRC64;			
QY	Query Match	100.0%; Score 31; DB 4; Length 95;			
QY	Best Local Similarity	100.0%; Pred. No. 99;			
QY	Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 SYWMS 5				
QY					
QY	30 SYWMS 34				
DB					

RESULT 2

Q92UT9 PRELIMINARY; PRT; 207 AA.

AC Q92UT9; 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Putative aldehyde or xanthine dehydrogenase, iron-sulfur subunit protein.
 GN R0999 OR SMB21558.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymb (megaplasmid 2).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Fuan T.W., Weidner S., Wong K., Buhmester J., Chain P.,
 RA Vorholster F.J., Hernandez-Lucas I., Becker A., Cowle A., Gouzy J.,
 RA Golding B., Puelher A.;
 RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL: AL603645; CAC9398.1; -;
 DR InterPro: IPR002888; 2Fe-2S bind.
 DR InterPro: IPR006058; 2Fe2S_Ferredoxin.
 DR InterPro: IPR010411; Ferredoxin.
 DR InterPro: IPR006311; Tat.
 DR Pfam: PF00111; fer2, 1.
 DR Pfam: PF01799; fer2, 2; 1.
 DR ProDom: PD166071; 2Fe-2S bind; 1.
 DR TIGRfams: TIGR01409; TAT_signal_seq; 1.
 DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
 DR KMW: Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 207 AA; 21523 MW; ADD4D86BF3D5FE7F CRC64;

Query Match 100.0%; Score 31; DB 16; Length 207;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
 Db 178 SYMMS 182

RESULT 3

Q8NRH2 PRELIMINARY; PRT; 395 AA.

ID Q8NRH2; 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE 2-poly(phenyl-6-methoxyphenol hydroxylase and related FAD-dependent
 DE oxidoreductases (EC 1.14.13.2).
 GN CG1077.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RT EMBL: AP005277; BAB98470.1; -;
 DR InterPro: IPR000733; Flav_monooxygenase.
 DR InterPro: IPR002938; MoxY_FAD_binding.
 DR InterPro: IPR003042; Rng_monooxygenase.
 DR Pfam: PF01494; FAD binding 3; 1.
 DR Pfam: PF01360; Monooxygenase; 1.

DR PRINTS; PR00420; ENGMONOXGNASE.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 395 AA; 44008 MW; C6F63B6946941C60 CRC64;

Query Match 100.0%; Score 31; DB 16; Length 395;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
 Db 343 SYMMS 347

RESULT 4

Q8FQ16 PRELIMINARY; PRT; 399 AA.

AC Q8FQ16; 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Putative 4-hydroxybenzoate 3-monooxygenase.
 GN CE1133.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
 RA Ileo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
 RA Ueda Y., Sugimoto S.;
 RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
 RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP005217; BAC17943.1; -;
 DR KMW: Monooxygenase; Complete proteome.
 SQ SEQUENCE 399 AA; 44698 MW; 6EFDDF41B4F30C14 CRC64;

Query Match 100.0%; Score 31; DB 16; Length 399;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
 Db 347 SYMMS 351

RESULT 5

Q8YOR3 PRELIMINARY; PRT; 939 AA.

ID Q8YOR3; 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein A1r3756.
 GN A1R3756.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasa moto S.,
 RA Matsumoto M., Itoh T., Yamagishi A., Nishio Y.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takasawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RT DNA Res. 8:205-213(2001).
 DR EMBL: AP003594; BAB75455.1; -;
 DR KMW: Hypothetical protein; Complete proteome.
 SQ SEQUENCE 939 AA; 104233 MW; 8FE0A7CA6C1759A5 CRC64;

Query Match 100.0%; Score 31; DB 16; Length 939;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
DB 931 SYWMS 935

RESULT 6

ID 013690 PRELIMINARY; PRT; 942 AA.

AC 013690;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative guanine nucleotide exchange factor.
GN SPAC11B3.11C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Murphy L., Harris D.;
RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Barrett B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; Z88595; CAB1190.1;-;
DR GeneDB; Spombe; SPAC11B3.11c; -;
DR InterPro; IPR001849; PH.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50190; SEC7; 1.
SQ SEQUENCE 942 AA; 105909 MW; A677CE2E619ECDF4 CRC64;

Query Match 100.0%; Score 31; DB 3; Length 942;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
DB 539 SYWMS 543

RESULT 7

ID 09SL18 PRELIMINARY; PRT; 942 AA.

AC 09SL18;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Mutator-like transposase.
GN AT2G05490.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiea.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.U., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.U., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana";
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC006220; AAD24658.1;-;
DR InterPro; IPR000566; Lipocin_cyFAPB.
DR InterPro; IPR004332; MDR.
DR Pfam; PF03108; MDR; 1.
DR Pfam; PF03108; MDR; 1.
DR SMART; SM00575; ZNF_PMZ; 1.
DR PROSITE; PS00213; LIPICININ; 1.
SQ SEQUENCE 942 AA; 106737 MW; 0DE83F5F5AC0B71E CRC64;

Query Match 100.0%; Score 31; DB 10; Length 942;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
DB 63 SYWMS 67

RESULT 8

ID 09S5K1 PRELIMINARY; PRT; 57 AA.

AC 09S5K1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I) (Fragment).
GN COI.
OS Mytilus trossulus (Blue mussel).
OC Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6551;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4;
RA Wares J.P., Cunningham C.W.;
RT "Comparative phylogeography of the North Atlantic.";
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2 AND COPPER B (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME C + 2 H(2)O.
CC -1- PATHWAY: RESPIRATORY CHAIN, TERMINAL STEP.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
DR EMBL; AF242032; AAK6961.1;-;
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR Copep; Electron transport; Heme; Inner membrane; Membrane; Oxidoreductase; Respiratory chain; Transmembrane; Transport;

KM Mitochondrion.
 FT NON_TER 1
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI.
 OS Mytilus trossulus (Blue mussel).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6551;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=759NF.
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY101434; AAM63559.1; -
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PRO1165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 1
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI.
 OS Mytilus edulis (Blue mussel).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=376WH;
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY101412; AAM63567.1; -
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PRO1165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 1
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI.
 OS Mytilus edulis (Blue mussel).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=366ME;
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-

DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI.
 OS Mytilus edulis (Blue mussel).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=376WH;
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY101412; AAM63567.1; -
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PRO1165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 1
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI.
 OS Mytilus edulis (Blue mussel).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=366ME;
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-

Query Match 90.3%; Score 28; DB 8; Length 105;
 Best Local Similarity 80.0%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 11

QY 1 SYMS 5
 DB 15 SYMS 19
 ID 08LXS3 PRELIMINARY; PRT; 107 AA.
 AC 08LXS3;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI.
 OS Mytilus edulis (Blue mussel).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=366ME;
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-

CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO 1 IS THE
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 CC AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 CC C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY101407; AAM63562.1; -.
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PRO1165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Respiratory chain; Transmembrane; Transport;
 FT NON_TER 1 1
 FT NON_TER 107 107
 SQ SEQUENCE 107 AA; 11598 MW; 5E6104D10C023051 CRC64;
 Query Match 90.3%; Score 28; DB 8; Length 107;
 Best Local Similarity 80.0%; Pred. No. 4e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYMS 5
 DB 17 SYMLS 21
 RESULT 12
 OBLXSD PRELIMINARY; PRT; 108 AA.
 ID OBLXSD:
 AC 08LXSD: 01-OCT-2002 (TREMblrel. 22, Created)
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI
 OS Mytilus edulis (Blue mussel).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A.NCI;
 RA Rignos C.; Hickerson M.J.; Henzler C.M.; Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 ATLantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO 1 IS THE
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 CC AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 CC C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY101420; AAM63575.1; -.
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PRO1165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport;
 KM Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 108 108

SQ SEQUENCE 108 AA; 11680 MW; 28D7C826D3297545 CRC64;
 Query Match 90.3%; Score 28; DB 8; Length 108;
 Best Local Similarity 80.0%; Pred. No. 4e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYMS 5
 DB 18 SYMLS 22
 RESULT 13
 OBLXSD PRELIMINARY; PRT; 108 AA.
 ID OBLXSD:
 AC 08LXSD: 01-OCT-2002 (TREMblrel. 22, Created)
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 GN COI
 OS Mytilus edulis (Blue mussel).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A.W10;
 RA Rignos C.; Hickerson M.J.; Henzler C.M.; Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 ATLantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO 1 IS THE
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 CC AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 CC C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY101417; AAM63572.1; -.
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PRO1165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport;
 KM Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 108 108
 SQ SEQUENCE 108 AA; 11712 MW; 28D7C826D3297545 CRC64;
 Query Match 90.3%; Score 28; DB 8; Length 108;
 Best Local Similarity 80.0%; Pred. No. 4e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYMS 5
 DB 18 SYMLS 22
 RESULT 14
 OBLXSD PRELIMINARY; PRT; 108 AA.
 ID OBLXSD:
 AC 08LXSD: 01-OCT-2002 (TREMblrel. 22, Created)
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Cytochrome oxidase subunit I (EC 1.9.3.1) (Cytochrome c oxidase

DE polypeptide 1) (Fragment).
 GN COI.
 OS Mytilus edulis (Blue mussel).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E.NC3.
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY130035; AAM83431.1; -.
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON TER 1 1
 FT NON TER 108 108
 SQ SEQUENCE 108 AA; 11762 MW; 68FC2E0C9337D118 CRC64;
 QY 1 SYNMS 5
 Db 26 SYMLS 30
 Query Match 90.3%; Score 28; DB 8; Length 108;
 Best Local Similarity 80.0%; Pred. No. 4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
 ID OBLXR9 PRELIMINARY; PRT; 108 AA.
 AC OBLXR9;
 DT 01-OCT-2002 (TREMBLER1.22, Created)
 DT 01-OCT-2002 (TREMBLER1.22, Last sequence update)
 DT 01-MAR-2003 (TREMBLER1.23, Last annotation update)
 DE Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)
 DE (Fragment).
 DE
 GN COI.
 OS Mytilus edulis (Blue mussel).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=G.FM99.1;
 RA Riggins C., Hickerson M.J., Henzler C.M., Cunningham C.W.;
 RT "A multilocus study of differential patterns of male and female trans-
 Atlantic gene flow in the blue mussel, Mytilus edulis.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2

CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 CC AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 C + 2 H(2)O.
 CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY101425; AAM63580.1; -.
 DR InterPro; IPR000883; COX1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON TER 1 1
 FT NON TER 108 108
 SQ SEQUENCE 108 AA; 11669 MW; 39C6DF8119537545 CRC64;
 QY 1 SYNMS 5
 Db 18 SYMLS 22
 Query Match 90.3%; Score 28; DB 8; Length 108;
 Best Local Similarity 80.0%; Pred. No. 4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Search completed: January 12, 2004, 07:01:54
 Job time : 15.1562 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Gen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 07:02:25 ; Search time 16 Seconds
(without alignments)
100.755 Million cell updates/sec

Title: US-09-829-495-66

Perfect score: 39

Sequence: 1 SYDSSNV 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	8	9	US-09-832-312-66
2	39	100.0	8	12	US-09-829-495-66
3	39	100.0	161	15	US-10-106-698-6785
4	33	84.6	617	15	US-10-205-342-11
5	32	82.1	166	11	US-09-797-464A-5
6	32	82.1	174	11	US-09-797-464A-9
7	32	82.1	365	11	US-09-797-464A-11
8	32	82.1	395	11	US-09-797-464A-2
9	32	82.1	893	14	US-10-014-436-4
10	32	82.1	1938	14	US-10-014-436-2
11	32	82.1	4327	12	US-10-369-493-10178
12	31	79.5	10	11	US-09-972-656-23
13	31	79.5	98	12	US-10-308-817-117
14	31	79.5	103	11	US-09-972-656-130
15	31	79.5	217	11	US-09-972-656-88

16	31	79.5	318	12	US-10-029-386-32107	Sequence 32107, A
17	31	79.5	595	15	US-10-074-475-257	Sequence 257, App
18	31	79.5	963	10	US-09-801-368-74	Sequence 74, App
19	31	79.5	963	12	US-10-369-493-1436	Sequence 1436, App
20	31	79.5	1420	12	US-10-369-493-11250	Sequence 11250, A
21	31	79.5	1610	15	US-10-155-533-9	Sequence 9, Appl1
22	30	76.9	241	11	US-09-805-354-14	Sequence 14, Appl1
23	30	76.9	241	15	US-10-144-259-14	Sequence 14, Appl1
24	30	76.9	242	11	US-09-805-354-15	Sequence 15, Appl1
25	30	76.9	242	15	US-10-144-259-15	Sequence 15, Appl1
26	30	76.9	341	12	US-10-240-145-146	Sequence 146, App
27	30	76.9	423	12	US-10-080-334-118	Sequence 118, App
28	30	76.9	428	12	US-10-080-334-117	Sequence 117, App
29	30	76.9	458	12	US-10-369-493-5399	Sequence 5399, App
30	30	76.9	596	12	US-10-080-334-14	Sequence 14, Appl1
31	30	76.9	618	9	US-09-765-272-72	Sequence 72, Appl1
32	30	76.9	627	12	US-10-262-794A-28	Sequence 28, Appl1
33	30	76.9	627	15	US-10-242-056-28	Sequence 28, Appl1
34	30	76.9	640	12	US-10-080-334-116	Sequence 116, Appl1
35	30	76.9	769	14	US-10-072-841-31	Sequence 31, Appl1
36	30	76.9	798	12	US-10-173-551-38	Sequence 38, Appl1
37	30	76.9	798	12	US-10-080-334-12	Sequence 12, Appl1
38	30	76.9	798	12	US-10-080-334-112	Sequence 112, App
39	30	76.9	805	12	US-10-080-334-115	Sequence 115, App
40	30	76.9	806	12	US-10-080-334-113	Sequence 113, App
41	30	76.9	806	12	US-10-080-334-114	Sequence 114, App
42	30	76.9	1189	12	US-10-262-794A-26	Sequence 26, Appl1
43	30	76.9	1190	15	US-10-242-056-26	Sequence 26, Appl1
44	29	74.4	129	16	US-10-174-693-307	Sequence 307, App
45	29	74.4	271	12	US-10-053-530-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-09-832-312-66
; Sequence 66, Application US/09832212
; Patent No. US20010049829A1
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-66
Query Match 100.0%; Score 39; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDSSNV 8
| | | | |
| | | | |
Db 1 SYDSSNV 8
US-09-829-495-66
; Sequence 66, Application US/09829495

Publication No. US20040001826A1
GENERAL INFORMATION:
APPLICANT: Busfield SJ
APPLICANT: Villevall J
APPLICANT: Jandrot-Perrus M
APPLICANT: Vainchenker W
APPLICANT: Gili DS
APPLICANT: Olan MD
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-234
CURRENT APPLICATION NUMBER: US/09/829,495
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 09/610,118
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/503,387
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 09/454,824
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 09/345,468
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PasteSeq for Windows Version 3.0
SEQ ID NO 66
LENGTH: 8
TYPE: PRT
ORGANISM: Homo sapiens
US-09-829-495-66

Query Match
Best Local Similarity 100.0%; Score 39; DB 12; Length 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
DB 1 SYDSSNVV 8

RESULT 3
US-10-106-698-6785
Sequence 6785, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PM005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 6785
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (149)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6785

Query Match
Best Local Similarity 100.0%; Score 39; DB 15; Length 161;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
DB 114 SYDSSNVV 121

RESULT 4
US-10-205-342-11
Sequence 11, Application US/10205342
Publication No. US20030108906A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alister
APPLICANT: Brooksbank, Robert
APPLICANT: Plimock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018198
CURRENT APPLICATION NUMBER: US/10/205,342
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
LENGTH: 617
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
OTHER INFORMATION: Protein: vacuolar adenosine triphosphatase subunit A
US-10-205-342-11

Query Match
Best Local Similarity 84.6%; Score 33; DB 15; Length 617;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
DB 184 NYDASNVV 191

RESULT 5
US-09-797-464A-5
Sequence 5, Application US/09797464A
Publication No. US20030022807A1
GENERAL INFORMATION:
APPLICANT: Wilting, Reinhard
APPLICANT: Bjornvad, Mads Eskelund
APPLICANT: Kauppinen, Markus Sakari
APPLICANT: Schuelein, Martin
TITLE OF INVENTION: Family 5 Xyloglucanases
FILE REFERENCE: 6073.200-US
CURRENT APPLICATION NUMBER: US/09/797,464A
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 166
TYPE: PRT
ORGANISM: Paenibacillus pabuli
US-09-797-464A-5

Query Match
Best Local Similarity 82.1%; Score 32; DB 11; Length 166;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
DB 108 SYDSSNVV 115

RESULT 6
US-09-797-464A-9
Sequence 9, Application US/09797464A
Publication No. US20030022807A1
GENERAL INFORMATION:
APPLICANT: Wilting, Reinhard
APPLICANT: Bjornvad, Mads Eskelund

```

; APPLICANT: Kauppinen, Markus Sakari
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: Family 5 Xyloglucanases
; FILE REFERENCE: 6073.200-US
; CURRENT APPLICATION NUMBER: US/09/797,464A
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Paenibacillus pabuli
US-09-797-464A-9

Query Match
Best Local Similarity 82.1%; Score 32; DB 11; Length 174;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
Db 117 SYDSSNVV 124

RESULT 7
US-09-797-464A-11
; Sequence 11, Application US/09797464A
; Publication No. US20030022807A1
; GENERAL INFORMATION:
; APPLICANT: Wiltling, Reinhard
; APPLICANT: Bjornvad, Mads Eskelund
; APPLICANT: Kauppinen, Markus Sakari
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: Family 5 Xyloglucanases
; FILE REFERENCE: 6073.200-US
; CURRENT APPLICATION NUMBER: US/09/797,464A
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Paenibacillus pabuli
US-09-797-464A-11

Query Match
Best Local Similarity 82.1%; Score 32; DB 11; Length 363;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
Db 299 SYDSSNVV 306

RESULT 8
US-09-797-464A-2
; Sequence 2, Application US/09797464A
; Publication No. US20030022807A1
; GENERAL INFORMATION:
; APPLICANT: Wiltling, Reinhard
; APPLICANT: Bjornvad, Mads Eskelund
; APPLICANT: Kauppinen, Markus Sakari
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: Family 5 Xyloglucanases
; FILE REFERENCE: 6073.200-US
; CURRENT APPLICATION NUMBER: US/09/797,464A
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Paenibacillus pabuli
US-09-797-464A-2
```

```

Query Match
Best Local Similarity 82.1%; Score 32; DB 11; Length 395;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
Db 331 SYDSSNVV 338

RESULT 9
US-10-014-436-4
; Sequence 4, Application US/10014436
; Publication No. US20020182699A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, Yuji
; APPLICANT: IGARASHI, Kazuaki
; APPLICANT: OZAKI, Katsuya
; APPLICANT: ARA, Katsutoshi
; APPLICANT: KAWAI, Shuji
; APPLICANT: ITO, Susumu
; TITLE OF INVENTION: Gene for Enzyme Having Both Alkaline Pullulanase and Alkaline Ali
; FILE REFERENCE: 2173-0122P
; CURRENT APPLICATION NUMBER: US/10/014,436
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: JP 111547/1995
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: PCT/JP96/01243
; PRIOR FILING DATE: 1996-05-10
; PRIOR APPLICATION NUMBER: US 08/952,084
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: US 09/514,302
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 893
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-10-014-436-4

Query Match
Best Local Similarity 82.1%; Score 32; DB 14; Length 893;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
Db 624 SYDSSDII 631

RESULT 10
US-10-014-436-2
; Sequence 2, Application US/10014436
; Publication No. US20020182699A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, Yuji
; APPLICANT: IGARASHI, Kazuaki
; APPLICANT: OZAKI, Katsuya
; APPLICANT: ARA, Katsutoshi
; APPLICANT: KAWAI, Shuji
; APPLICANT: ITO, Susumu
; TITLE OF INVENTION: Gene for Enzyme Having Both Alkaline Pullulanase and Alkaline Ali
; FILE REFERENCE: 2173-0122P
; CURRENT APPLICATION NUMBER: US/10/014,436
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: JP 111547/1995
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: PCT/JP96/01243
; PRIOR FILING DATE: 1996-05-10
; PRIOR APPLICATION NUMBER: US 08/952,084
; PRIOR FILING DATE: 1997-11-10
```

```
/ PRIOR APPLICATION NUMBER: US 09/514,302
/ PRIOR FILING DATE: 2000-02-28
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 1938
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
US-10-014-436-2

Query Match      82.1%; Score 32; DB 14; Length 1938;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYDSSNV 8
Db      1669 SYDSDII 1676

RESULT 11
US-10-369-493-10178
/ Sequence 10178, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 10178
/ LENGTH: 4327
/ TYPE: PRT
/ ORGANISM: magnetite-containing magnetic coccus
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(4327)
/ OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-10178

Query Match      82.1%; Score 32; DB 12; Length 4327;
Best Local Similarity 85.7%; Pred. No. 2.7e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYDSSNV 7
Db      3732 AYDSSNV 3738

RESULT 12
US-09-972-656-23
/ Sequence 23, Application US/09972656
/ Publication No. US20030099647A1
/ GENERAL INFORMATION:
/ APPLICANT: Deshpande, Rajendra
/ APPLICANT: Tsai, Mei-Mei
/ TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
/ FILE REFERENCE: A-799
/ CURRENT APPLICATION NUMBER: US/09/972,656
/ CURRENT FILING DATE: 2001-10-05
/ NUMBER OF SEQ ID NOS: 135
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 23
/ LENGTH: 10
/ TYPE: PRT
```

```
/ ORGANISM: Homo sapiens
US-09-972-656-23

Query Match      79.5%; Score 31; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYDSSN 6
Db      2 SYDSSN 7

RESULT 13
US-10-308-817-117
/ Sequence 117, Application US/10308817
/ Publication No. US20030219661A1
/ GENERAL INFORMATION:
/ APPLICANT: Rother, Russell
/ APPLICANT: Wu, Dayang
/ TITLE OF INVENTION: HYBRID ANTIBODIES
/ FILE REFERENCE: 1087-37
/ CURRENT APPLICATION NUMBER: US/10/308,817
/ CURRENT FILING DATE: 2002-12-03
/ NUMBER OF SEQ ID NOS: 195
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 117
/ LENGTH: 98
/ TYPE: PRT
/ ORGANISM: human
US-10-308-817-117

Query Match      79.5%; Score 31; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYDSSN 6
Db      93 SYDSSN 98

RESULT 14
US-09-972-656-130
/ Sequence 130, Application US/09972656
/ Publication No. US20030099647A1
/ GENERAL INFORMATION:
/ APPLICANT: Deshpande, Rajendra
/ APPLICANT: Tsai, Mei-Mei
/ TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
/ FILE REFERENCE: A-799
/ CURRENT APPLICATION NUMBER: US/09/972,656
/ CURRENT FILING DATE: 2001-10-05
/ NUMBER OF SEQ ID NOS: 135
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 130
/ LENGTH: 103
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: Misc.
/ LOCATION: (7)..(7)
/ OTHER INFORMATION: unidentified
/ NAME/KEY: Misc.
/ LOCATION: (33)..(33)
/ OTHER INFORMATION: unidentified
/ NAME/KEY: Misc.
/ LOCATION: (34)..(34)
/ OTHER INFORMATION: unidentified
/ NAME/KEY: Misc.
/ LOCATION: (35)..(35)
/ OTHER INFORMATION: unidentified
/ NAME/KEY: Misc.
/ LOCATION: (36)..(36)
```

; OTHER INFORMATION: Unidentified
US-09-972-656-130

Query Match 79.5%; Score 31; DB 11; Length 103;
Best Local Similarity 100.0%; Pred No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYDSSN 6
|||
Db 98 SYDSSN 103

RESULT 15

US-09-972-656-88
; Sequence 88, Application US/09972656
; Publication No. US2003009647A1
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 88
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-88

Query Match 79.5%; Score 31; DB 11; Length 217;
Best Local Similarity 100.0%; Pred No. 188+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYDSSN 6
|||
Db 93 SYDSSN 98

Search completed: January 12, 2004, 07:19:58
Job time : 18 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:56:19 ; Search time 6.625 Seconds
(without alignments)
51.092 Million cell updates/sec

Title: US-09-829-495-66

Perfect score: 39

Sequence: 1 SYDSSNV 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	8	4	US-09-832-312-66
2	32	82.1	515	4	US-09-107-532A-5317
3	32	82.1	893	4	US-09-514-302-4
4	32	82.1	1938	4	US-09-514-302-2
5	31	79.5	222	1	US-07-869-933-12
6	31	79.5	222	1	US-07-869-933-28
7	31	79.5	222	3	US-09-103-663-12
8	31	79.5	222	3	US-09-103-663-28
9	31	79.5	222	3	US-08-311-731A-45
10	30	76.9	126	4	US-09-485-973-6
11	30	76.9	252	5	PCT-US96-01314-58
12	30	76.9	253	5	PCT-US96-01314-53
13	30	76.9	618	4	US-08-961-083-72
14	30	76.9	618	4	US-09-536-784-72
15	30	76.9	627	4	US-08-851-567B-28
16	30	76.9	676	3	US-08-630-172-10
17	30	76.9	676	3	US-09-375-419-10
18	30	76.9	769	2	US-08-789-078-1
19	30	76.9	769	2	US-08-752-633-1
20	30	76.9	769	2	US-08-476-062A-45
21	30	76.9	769	2	US-07-728-215-31
22	30	76.9	769	4	US-08-938-085A-31
23	30	76.9	769	4	US-10-072-844-31
24	30	76.9	769	5	PCT-US95-04886-1
25	30	76.9	769	5	PCT-US96-01314-45
26	30	76.9	796	4	US-09-107-532A-7065
27	30	76.9	891	4	US-09-328-352-6637

28	30	76.9	1189	4	US-08-851-567B-26	Sequence 26, Appl
29	29	74.4	16	2	US-08-932-876-1	Sequence 1, Appl
30	29	74.4	25	1	US-08-378-761A-30	Sequence 30, Appl
31	29	74.4	25	1	US-08-485-286-30	Sequence 30, Appl
32	29	74.4	25	6	5248606-16	Parent No. 5248606
33	29	74.4	40	2	US-08-679-405-6	Sequence 6, Appl
34	29	74.4	40	2	US-08-842-799-6	Sequence 6, Appl
35	29	74.4	40	5	PCT-US96-11458-6	Sequence 6, Appl
36	29	74.4	109	3	US-09-157-370-5	Sequence 307, App
37	29	74.4	129	4	US-09-615-192A-307	Sequence 307, App
38	29	74.4	357	4	US-09-252-991A-24337	Sequence 24337, A
39	29	74.4	469	3	US-08-378-313-33	Sequence 33, Appl
40	29	74.4	599	4	US-09-832-441-2	Sequence 2, Appl
41	29	74.4	599	4	US-09-833-102-2	Sequence 2, Appl
42	29	74.4	600	2	US-08-679-405-2	Sequence 2, Appl
43	29	74.4	600	3	US-08-842-799-2	Sequence 2, Appl
44	29	74.4	600	3	US-09-271-778-2	Sequence 2, Appl
45	29	74.4	600	4	US-09-788-871-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-09-832-312-66
; Sequence 66, Application US/09832312
; Patent No. 6548741
; GENERAL INFORMATION:
; APPLICANT: Busfield et al
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-66

Query Match
Best Local Similarity 100.0%; Score 39; DB 4; Length 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
DB 1 SYDSSNV 8

RESULT 2
US-09-107-532A-5317
; Sequence 5317, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA

```

ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,488
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8877
INFORMATION FOR SEQ ID NO: 5317:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...515
SEQUENCE DESCRIPTION: SEQ ID NO: 5317:
US-09-107-532A-5317

Query Match 82.1%; Score 32; DB 4; Length 515;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSNV 7
:|||||
DB 175 AYDSNV 181

RESULT 3
US-09-514-302-4
Sequence 4, Application US/09514302
Patent No. 6338959
GENERAL INFORMATION:
APPLICANT: HATADA, Yuji
APPLICANT: IGARASHI, Kazuaki
APPLICANT: OZAKI, Katsuya
APPLICANT: ARA, Katsutoshi
APPLICANT: KAWAI, Shuji
APPLICANT: ITO, Susumu
TITLE OF INVENTION: GENE FOR ENZYME HAVING BOTH ALKALINE PHOSPHATASE AND
TITLE OF INVENTION: ALKALINE PHOSPHATASE ACTIVITIES
FILE REFERENCE: 2173-105P
CURRENT APPLICATION NUMBER: US/09/514,302
CURRENT FILING DATE: 2000-02-28
EARLIER APPLICATION NUMBER: 08/952,084
EARLIER FILING DATE: 1997-11-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 893
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-514-302-4

Query Match 82.1%; Score 32; DB 4; Length 893;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDSNV 8
:|||||
DB 624 SYDSNV 631

RESULT 4
US-09-514-302-2
Sequence 2, Application US/09514302
Patent No. 6338959
GENERAL INFORMATION:
APPLICANT: HATADA, Yuji
APPLICANT: IGARASHI, Kazuaki
APPLICANT: OZAKI, Katsuya
APPLICANT: ARA, Katsutoshi
APPLICANT: KAWAI, Shuji
APPLICANT: ITO, Susumu
TITLE OF INVENTION: GENE FOR ENZYME HAVING BOTH ALKALINE PHOSPHATASE AND
TITLE OF INVENTION: ALKALINE PHOSPHATASE ACTIVITIES
FILE REFERENCE: 2173-105P
CURRENT APPLICATION NUMBER: US/09/514,302
CURRENT FILING DATE: 2000-02-28
EARLIER APPLICATION NUMBER: 08/952,084
EARLIER FILING DATE: 1997-11-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 1938
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-514-302-2

Query Match 82.1%; Score 32; DB 4; Length 1938;
Best Local Similarity 62.5%; Pred. No. 6.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSNV 8
:|||||
DB 1669 SYDSNV 1676

RESULT 5
US-07-869-933-12
Sequence 12, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIH
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Rat
STRAIN: PCRI alpha subunit
US-07-869-933-12

Query Match 79.5%; Score 31; DB 1; Length 222;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
Db 130 SYDSSNNI 136

RESULT 6
US-07-869-933-28
Sequence 28, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
STRAIN: alpha subunit
US-07-869-933-28

Query Match 79.5%; Score 31; DB 1; Length 222;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
Db 130 SYDSSNNI 136

Db 130 SYDSSNNI 136

RESULT 7
US-09-103-663-12
Sequence 12, Application US/09103663D
Patent No. 6171803
GENERAL INFORMATION:
APPLICANT: Kinet et al.
TITLE OF INVENTION: Isolation, characterization, and use of the human beta
TITLE OF INVENTION: immunoglobulin E.
FILE REFERENCE: 50490
CURRENT APPLICATION NUMBER: US/09/103,663D
CURRENT FILING DATE: 1998-06-23
EARLIER APPLICATION NUMBER: 07/869,933
EARLIER FILING DATE: 1992-04-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 222
TYPE: PRT
ORGANISM: Rattus sp.
US-09-103-663-12

Query Match 79.5%; Score 31; DB 3; Length 222;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
Db 130 SYDSSNNI 136

RESULT 8
US-09-103-663-28
Sequence 28, Application US/09103663D
Patent No. 6171803
GENERAL INFORMATION:
APPLICANT: Kinet et al.
TITLE OF INVENTION: Isolation, characterization, and use of the human beta
TITLE OF INVENTION: subunit of the high affinity receptor for
FILE REFERENCE: 50490
CURRENT APPLICATION NUMBER: US/09/103,663D
CURRENT FILING DATE: 1998-06-23
EARLIER APPLICATION NUMBER: 07/869,933
EARLIER FILING DATE: 1992-04-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapiens
US-09-103-663-28

Query Match 79.5%; Score 31; DB 3; Length 222;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
Db 130 SYDSSNNI 136

RESULT 9
US-08-311-731A-45
Sequence 45, Application US/08311731A
Patent No. 6583266
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-45

Query Match 79.5%; Score 31; DB 4; Length 382;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSNVV 8
DB 147 AYDSPNVV 154

RESULT 10
US-09-485-973-6
Sequence 6, Application US/09485973
Patent No. 6488935
GENERAL INFORMATION:
APPLICANT: De Villiers Zur Hausen, Ethel-Micelle
APPLICANT: Zur Hausen, Harald
TITLE OF INVENTION: PAPILLOMA VIRUS, AGENTS FOR THE DETECTION THEREOF AND FOR THE
TITLE OF INVENTION: THERAPY
TITLE OF INVENTION: OF THE DISEASES CAUSED BY SAID VIRUS
FILE REFERENCE: 008484-0082-999
CURRENT APPLICATION NUMBER: US/09/485,973
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: PCT/DE98/02379
PRIOR FILING DATE: 1998-08-12
PRIOR APPLICATION NUMBER: 197 35 118.2
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin version 3.0
SEQ ID NO 6
LENGTH: 126
TYPE: PRT
ORGANISM: Papillomavirus Capsid Protein
US-09-485-973-6

Query Match 76.9%; Score 30; DB 4; Length 126;
Best Local Similarity 83.3%; Pred. No. 76;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNVV 7
DB 30 YDSSNI 35

RESULT 11
PCT-US96-01314-58
Sequence 58, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-58

Query Match 76.9%; Score 30; DB 5; Length 252;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSNVV 8
DB 237 SEDSNV 244

RESULT 12
PCT-US96-01314-53
Sequence 53, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.

ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-53

Query Match 76.9%; Score 30; DB 5; Length 253;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SYDSSNV 8
Db 238 SYDSSNV 245

RESULT 13
US-08-961-083-72
Sequence 72, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 72:

SEQUENCE CHARACTERISTICS:
LENGTH: 618 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-72

Query Match 76.9%; Score 30; DB 3; Length 618;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SYDSSNV 8
Db 363 SYDSSDAV 370

RESULT 14
US-09-536-784-72
Sequence 72, Application US/09536784
Patent No. 6573082
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,784
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 618 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-536-784-72

Query Match 76.9%; Score 30; DB 4; Length 618;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SYDSSNV 8
Db 363 SYDSSDAV 370

RESULT 15
US-08-851-567B-28
Sequence 28, Application US/08851567B

Search completed: January 12, 2004, 07:04:03
Job time : 7.625 secs

Patent No. 6528484
GENERAL INFORMATION:
APPLICANT: Ensign, Jerald C
APPLICANT: Bowen, David J
APPLICANT: Petrell, James
APPLICANT: Fatig, Raymond
APPLICANT: Schoonover, Sue
APPLICANT: French-Constant, Richard
APPLICANT: Rocheleau, Thomas A.
APPLICANT: Blackburn, Michael B.
APPLICANT: Hey, Timothy D.
APPLICANT: Merlo, Donald J.
APPLICANT: Orr, Gregory L.
APPLICANT: Robertes, Jean L.
APPLICANT: Strickland, James A.
APPLICANT: Guo, Lining
APPLICANT: Cliche, Todd A.
APPLICANT: Sukhapinda, Kitiseri
TITLE OF INVENTION: Insecticidal Protein Toxins From Photobadus
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: IN
COUNTRY: US
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,567B
FILING DATE: 05-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/063,615
FILING DATE: 18-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/395,497
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,255
FILING DATE: 06-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,423
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/705,484
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.93804
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-851-567B-28

Query Match 76.9%; Score 30; DB 4; Length 627;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 YDSSNVV 8
DB 336 YDSGNVI 342

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:56:19 ; Search time 5.79688 Seconds
(without alignments)
51.092 Million cell updates/sec

Title: US-09-829-495-65
Perfect score: 38
Sequence: 1 EDNORPS 7

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	7	4	US-09-832-312-65
2	34	89.5	106	1	US-08-259-372A-16
3	34	89.5	106	1	US-08-468-671-16
4	33	86.8	112	2	US-08-665-202-39
5	33	86.8	112	2	US-09-315-574-39
6	32	84.2	108	1	US-08-259-372A-10
7	32	84.2	108	1	US-08-468-671-10
8	32	84.2	108	1	US-08-468-671-10
9	31	81.6	64	2	US-09-252-991A-27287
10	31	81.6	108	4	US-08-765-179B-19
11	31	81.6	109	3	US-09-025-769B-20
12	31	81.6	249	4	US-10-039-785-53
13	30	78.9	109	4	US-09-025-769B-32
14	30	78.9	109	4	US-09-025-769B-51
15	30	78.9	234	4	US-09-372-425A-4
16	30	78.9	462	3	US-08-875-944B-5
17	30	78.9	462	3	US-09-116-049-4
18	30	78.9	462	4	US-09-602-868A-5
19	30	78.9	462	4	US-09-884-563-4
20	30	78.9	555	2	US-08-780-835B-2
21	30	78.9	555	3	US-09-303-268-2
22	30	78.9	555	3	US-09-116-049-2
23	30	78.9	555	4	US-09-884-563-2
24	30	78.9	620	4	US-08-426-509A-9
25	30	78.9	620	4	US-08-232-545-9
26	30	78.9	620	5	PCT-US95-05008-9
27	29	76.3	109	1	US-08-478-039-91

28	29	76.3	109	1	US-08-476-349A-91	Sequence 91, Appl
29	29	76.3	111	3	US-08-983-607-35	Sequence 35, Appl
30	29	76.3	112	3	US-08-983-607-31	Sequence 31, Appl
31	29	76.3	132	1	US-08-534-975-4	Sequence 4, Appl
32	29	76.3	132	2	US-08-954-470-4	Sequence 4, Appl
33	29	76.3	132	3	US-09-129-855A-4	Sequence 4, Appl
34	29	76.3	132	3	US-09-247-154-4	Sequence 4, Appl
35	29	76.3	132	4	US-09-480-718-4	Sequence 4, Appl
36	29	76.3	132	4	US-09-610-833-4	Sequence 4, Appl
37	29	76.3	132	4	US-09-129-855A-4	Sequence 4, Appl
38	29	76.3	139	4	US-09-252-991A-19245	Sequence 19245, A
39	29	76.3	475	4	US-09-252-991A-29248	Sequence 29248, A
40	29	76.3	743	4	US-09-328-352-5073	Sequence 5073, Ap
41	28	73.7	7	1	US-08-264-093-19	Sequence 19, Appl
42	28	73.7	59	4	US-09-638-715-33	Sequence 33, Appl
43	28	73.7	93	4	US-10-060-509-33	Sequence 33, Appl
44	28	73.7	93	4	US-08-936-165A-466	Sequence 466, App
45	28	73.7	98	1	US-08-211-202-111	Sequence 111, App

ALIGNMENTS

```

RESULT 1
US-09-832-312-65
; Sequence 65, Application US/09832312
; Patent No. 6548741
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 65
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-65

Query Match      100.0%; Score 38; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDNORPS 7
      |||||
Db      1 EDNORPS 7

RESULT 2
US-08-259-372A-16
; Sequence 16, Application US/08259372A
; Patent No. 5565354
; GENERAL INFORMATION:
; APPLICANT: Ostberg, Lars G.
; TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA

```

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,372A
FILING DATE: 14-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/871,426
FILING DATE: 21-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,036
FILING DATE: 27-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/538,796
FILING DATE: 15-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/192,754
FILING DATE: 11-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/925,196
FILING DATE: 31-OCT-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/904,517
FILING DATE: 05-SEP-1986
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-50-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-259-372A-16

Query Match 89.5%; Score 34; DB 1; Length 106;
Best Local Similarity 85.7%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDNORPS 7
Db 49 EDNKRPS 55

RESULT 3
US-08-468-671-16
Sequence 16, Application US/08468671
Patent No. 5648077
GENERAL INFORMATION:
APPLICANT: Ostberg, Lars G.
TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,671
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,372
FILING DATE: 14-JUN-1994
APPLICATION NUMBER: US 07/871,426
FILING DATE: 21-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,036
FILING DATE: 27-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/538,796
FILING DATE: 15-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/192,754
FILING DATE: 11-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/925,196
FILING DATE: 31-OCT-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/904,517
FILING DATE: 05-SEP-1986
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-50-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-671-16

Query Match 89.5%; Score 34; DB 1; Length 106;
Best Local Similarity 85.7%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDNORPS 7
Db 49 EDNKRPS 55

RESULT 4
US-08-665-202-39
Sequence 39, Application US/08665202
Patent No. 597322
GENERAL INFORMATION:
APPLICANT: Marks, James D.
TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,202
FILING DATE: 13-JUN-1996

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061410
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-665-202-39

Query Match 86.8%; Score 33; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DNORPS 7
|||||
Db 52 DNORPS 57

RESULT 5
US-09-315-574-39
Sequence 39, Application US/09315574
Patent No. 6512097
GENERAL INFORMATION:
APPLICANT: Marks, James D.
TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Majestic, Parsons, Siebert & Haue P.C.
STREET: Four Embarcadero Center, Suite 1100
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,574
FILING DATE: 20-MAY-99
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061411
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-315-574-39

Query Match 86.8%; Score 33; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DNORPS 7
|||||
Db 52 DNORPS 57

RESULT 6
US-08-259-372A-10
Sequence 10, Application US/08259372A
Patent No. 5565354
GENERAL INFORMATION:
APPLICANT: Ostberg, Lars G.
TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,372A
FILING DATE: 14-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/871,426
FILING DATE: 21-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,036
FILING DATE: 27-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/538,796
FILING DATE: 15-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/192,754
FILING DATE: 11-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/925,196
FILING DATE: 31-OCT-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/904,517
FILING DATE: 05-SEP-1986
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-50-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-259-372A-10

Query Match 84.2%; Score 32; DB 1; Length 108;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 49 DDNORPS 55

RESULT 7
US-08-468-671-10
Sequence 10, Application US/08468671
Patent No. 5648077
GENERAL INFORMATION:
APPLICANT: Oseberg, Lars G.
TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,671
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,372
FILING DATE: 14-JUN-1994
APPLICATION NUMBER: US 07/871,426
FILING DATE: 21-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,036
FILING DATE: 27-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/538,796
FILING DATE: 15-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/192,754
FILING DATE: 11-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/925,196
FILING DATE: 31-OCT-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/904,517
FILING DATE: 05-SEP-1986
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-50-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-671-10

Query Match 84.2%; Score 32; DB 1; Length 108;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 49 DDNORPS 55

RESULT 8
US-09-252-991A-27287
Sequence 27287, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27287
LENGTH: 192
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27287

Query Match 84.2%; Score 32; DB 4; Length 192;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 4 EDNORPS 10

RESULT 9
US-08-765-179B-19
Sequence 19, Application US/08765179B
Patent No. 5854027
GENERAL INFORMATION:
APPLICANT: STEIPE, Boris
APPLICANT: STEINBACHER, Stefan
TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY
TITLE OF INVENTION: OF ANTIBODIES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,179B
FILING DATE: 14-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02626
FILING DATE: 06-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 44 25 115.7
FILING DATE: 15-JUL-1994
INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 64 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-179B-19

Query Match 81.6%; Score 31; DB 2; Length 64;
Best Local Similarity 71.4%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
:|||||
DB 19 DDKRPS 25

RESULT 10
US-09-025-769B-20

Sequence 20, Application US/09025769B
Patent No. 630064

GENERAL INFORMATION:

APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Ilaq, Vic
APPLICANT: Ge, Jaming
APPLICANT: Moroney, Simon
APPLICANT: Plueckhuhn, Andreas
TITLE OF INVENTION: Protein(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 95 11 3021.0

FILING DATE: 18-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: James F. Haley, Jr., Esq.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: MORPHO/5

TELEPHONE: (212)596-9000

TELEFAX: (212)596-9090

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 108 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-025-769B-20

Query Match 81.6%; Score 31; DB 4; Length 108;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
:|||||
DB 48 DDKRPS 54

RESULT 11

US-09-157-370-5

Sequence 5, Application US/09157370A

Patent No. 626238

GENERAL INFORMATION:

APPLICANT: STEIBE, Boris

APPLICANT: STEINBACHER, Stefan

TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY OF ANTIBODIES

FILE REFERENCE: P8341-8072

CURRENT APPLICATION NUMBER: US/09/157,370A

CURRENT FILING DATE: 1998-09-21

EARLIER APPLICATION NUMBER: 08/765,179

EARLIER FILING DATE: 1997-01-14

EARLIER APPLICATION NUMBER: PCT/EP95/02626

EARLIER FILING DATE: 1995-07-06

EARLIER APPLICATION NUMBER: DE/P44 25 115.7

EARLIER FILING DATE: 1994-07-15

NUMBER OF SEQ ID NOS: 10

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 5

LENGTH: 109

TYPE: PRT

ORGANISM: Homo sapiens

US-09-157-370-5

QY 1 EDNRPS 7
:|||||
DB 49 DDKRPS 55

Query Match 81.6%; Score 31; DB 3; Length 109;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
US-10-039-785-53

Sequence 53, Application US/10039785

Patent No. 6538938

GENERAL INFORMATION:

APPLICANT: Salcedo et al.

TITLE OF INVENTION: Antibodies that Immunospecifically Bind to TRAIL

FILE REFERENCE: PF550

CURRENT APPLICATION NUMBER: US/10/039,785

CURRENT FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 60/369,860

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: 60/341,237

PRIOR FILING DATE: 2001-12-20

PRIOR APPLICATION NUMBER: 60/331,310

PRIOR FILING DATE: 2001-11-14

PRIOR APPLICATION NUMBER: 60/331,044

PRIOR FILING DATE: 2001-11-07

PRIOR APPLICATION NUMBER: 60/327,364

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/323,807

PRIOR FILING DATE: 2001-09-21

PRIOR APPLICATION NUMBER: 60/309,176

PRIOR FILING DATE: 2001-08-02

PRIOR APPLICATION NUMBER: 60/294,981

PRIOR FILING DATE: 2001-06-04

PRIOR APPLICATION NUMBER: 60/293,473

PRIOR FILING DATE: 2001-05-25

NUMBER OF SEQ ID NOS: 66

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 53

LENGTH: 249

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: T1006F07 scFv

US-10-039-785-53

Query Match 81.6%; Score 31; DB 4; Length 249;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
:|||||
Db 190 QDNKRPS 196

RESULT 13

US-09-025-769B-32
; Sequence 32, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-32

Query Match 78.9%; Score 30; DB 4; Length 109;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
:|||||
Db 51 DNNRPS 57

RESULT 14
US-09-025-769B-51
; Sequence 51, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic

APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-51

Query Match 78.9%; Score 30; DB 4; Length 109;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
:|||||
Db 51 DNNRPS 57

RESULT 15
US-09-372-425A-4
; Sequence 4, Application US/09372425A
; Patent No. 6475749
; GENERAL INFORMATION:
; APPLICANT: Sherie L. Morrison
; APPLICANT: Ramon Montano
; TITLE OF INVENTION: Improved Rh Antibody
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 98
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/372,425A
; FILING DATE: August 11, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 510015-223
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 788-5100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 234 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Light chain - AA
US-09-372-425A-4

Query Match 78.9%; Score 30; DB 4; Length 234;
Best Local Similarity 71.4%; Pred. No. 94;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNQRP 7
:::|||||
Db 69 DNNQRP 75

Search completed: January 12, 2004, 07:04:02
Job time : 6.79668 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2004, 06:54:49 ; Search time 7.25 Seconds
(without alignments)
106.117 Million cell updates/sec

Title: US-09-829-495-66

Perfect score: 39
Sequence: 1 SYDSSNVV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	505	2 AC1469	internalin like pr
2	33	84.6	680	2 T08080	probable myrosinase
3	33	84.6	730	2 P96559	hypothetical prote
4	33	84.6	934	2 C81298	probable formate d
5	33	84.6	1037	2 A60163	glycoprotein fib
6	33	84.6	1945	2 T13937	plexin A - fruit f
7	32	82.1	112	1 L6H48	ig lambda chain v-
8	32	82.1	469	2 B37837	probable alpha-amy
9	32	82.1	527	2 E88040	protein F47F6.3 [l
10	32	82.1	571	2 S58356	pept protein - Sta
11	31	79.5	130	2 T20762	hypothetical prote
12	31	79.5	245	2 A30154	IGF receptor alpha
13	31	79.5	261	2 T15520	hypothetical prote
14	31	79.5	271	2 UC4584	insulin-like growt
15	31	79.5	304	2 H90095	hypothetical prote
16	31	79.5	358	2 C96923	similar to chlorom
17	31	79.5	374	2 T21513	hypothetical prote
18	31	79.5	411	2 B87200	probable ligase [l
19	31	79.5	413	2 A70795	hypothetical prote
20	31	79.5	450	2 T21515	hypothetical prote
21	31	79.5	783	2 B84823	probable isoamylas
22	31	79.5	876	2 B96693	probable receptor
23	31	79.5	963	2 S45167	chitin synthase [E
24	30	76.9	229	2 S49494	replication protei
25	30	76.9	232	2 AH0954	probable membrane
26	30	76.9	232	2 T23806	hypothetical prote
27	30	76.9	311	2 T17613	hypothetical prote
28	30	76.9	344	2 H97223	hypothetical prote
29	30	76.9	344	2 H97223	hypothetical prote

30	30	76.9	358	2 T15069	hypothetical prote
31	30	76.9	429	2 S63360	hypothetical prote
32	30	76.9	458	2 T16041	hypothetical prote
33	30	76.9	462	2 T52115	myrosinase-binding
34	30	76.9	462	2 B96560	hypothetical prote
35	30	76.9	624	1 VCNV87	p87 capsid protein
36	30	76.9	624	2 T10374	p87 capsid protein
37	30	76.9	625	2 A84787	hypothetical prote
38	30	76.9	630	2 JQ1894	hypothetical 74.1K
39	30	76.9	646	2 T02398	hypothetical prote
40	30	76.9	754	2 S60464	mes-3 protein - Ca
41	30	76.9	754	2 G87767	mes-3 protein - Ca
42	30	76.9	769	1 IJHULM	leukocyte adhesion
43	30	76.9	798	2 A40526	integrin beta-7 ch
44	30	76.9	803	2 B84931	DNA topoisomerase
45	30	76.9	806	2 A46271	integrin beta-7 ch

ALIGNMENTS

RESULT 1
AC1469
internalin like protein (LPXTG motif) [imported] - Listeria innocua (strain Clijp11262)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AC1469
R/Glaier, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baguer, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fath, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tietze, A.; Vazquez-Roland, J.A.; Voos, H.; Wehland,
A./Title: Comparative genomics of Listeria species.
A/Reference number: ABI077; PMID:21537279; PMID:11679665
A/Accession: AC1469
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-505 <GLA>
A/Cross-references: GB:AL592022; PIRN:CA95523.1; PID:g16412719; GSPDB:GN00178
A/Experimental source: strain Clijp11262
C/Genetics:
A/Gene: lln0290

Query Match
Best Local Similarity 84.6%; Score 33; DB 2; Length 505;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY : ||| |||
DB 319 TYDSSNVV 326

RESULT 2
T08080
probable myrosinase-binding protein - rape
N/Alternate names: jasmonate inducible protein
C/Species: Brassica napus (rape)
C/Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C/Accession: T08080
R/Gesht, N.; Brandt, A.
Planta 204, 295-304, 1998
A/Title: Two jasmonate-inducible myrosinase-binding proteins from Brassica napus L. seed
A/Reference number: Z16340; PMID:98192006; PMID:9530873
A/Accession: T08080
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-680 <GEN>
A/Cross-references: EMBL:Y11483; NID:e1023101; PIRN:CAA72271.1; PID:c304150
A/Experimental source: cv. Global; isolate a4; young seedlings
A/Note: jasmonate inducible

Query Match 84.6%; Score 33; DB 2; Length 680;

Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNV 8
DB 308 YDSSNII 314

RESULT 3

P96559 hypothetical protein F5P19.6 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: P96559

R/Rheology: A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: P96559

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-730 <STO>

A/Cross-references: GB:AE005173; NID:94220464; PIDN:AAD12691.1; GSPDB:GND0141

C/Genetics:

A:Gene: F5P19.6

A:Map position: 1

Query Match 84.6%; Score 33; DB 2; Length 730;
Best Local Similarity 71.4%; Pred. No. 56;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNV 8
DB 511 YDSSNII 517

RESULT 4

C81298 probable formate dehydrogenase (EC 1.2.1.2) large chain Cj151c [similarity] - Campylobacter

C/Species: Campylobacter jejuni

C/Date: 31-Mar-2000 #sequence_revision 15-Sep-2000 #text_change 03-Jun-2002

C/Accession: C81298

R/Parhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVlier, A.; Whitehead, S.; Barrel Nature 403, 665-668, 2000

A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A/Reference number: A81250; MUID:20150912; PMID:10688204

A/Accession: C81298

A/Molecule type: DNA

A/Residues: 1934 <PAR>

A/Cross-references: GB:AL119078; GB:AL111168; NID:96968723; PIDN:CBW73932.1; PID:9696893

A/Experimental source: serotype O2, strain NCTC 11168

A/Note: a translation exception was noted by the authors

A/Note: in Genbank entry AL119078, release 117.0, PIDN:CBW73932.1, the selenocysteine U

C/Genetics:

A:Gene: fdhA, Cj1511c

C/Superfamily: formate dehydrogenase

C/Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein; molybdenum; molybdopterin; NAD;

F;57/60/64/92/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

F;94/182/Active site: lys, His #status predicted

F;181/Binding site: molybdopterin guanine dinucleotide (Cys) (covalent) #status predicted

F;181/Modified site: selenocysteine #status predicted

F;181/Modified site: selenocysteine #status predicted

Query Match 84.6%; Score 33; DB 2; Length 934;
Best Local Similarity 62.5%; Pred. No. 73;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
DB 714 SYDSSNII 721

RESULT 5

A60163 glycoprotein IIB - rat

C/Species: Rattus sp. (rat)

C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 13-Sep-1998

C/Accession: A60163; B60163

R/Poncz, M.; Newman, P.J.

Blood 75, 1282-1289, 1990

A/Title: Analysis of rodent platelet glycoprotein IIB: evidence for evolutionarily conse

A/Reference number: A60163; MUID:90181615; PMID:2310828

A/Accession: A60163

A/Status: nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-1037 <PON>

A/Accession: B60163

A/Molecule type: protein

A/Residues: 32-50 <PON>

C/Comment: this protein is proteolytically processed into a heavy chain and a light cha

C/Superfamily: integrin alpha-2b chain

C/Keywords: calcium; disulfide bond; glycoprotein

Query Match 84.6%; Score 33; DB 2; Length 1037;
Best Local Similarity 87.5%; Pred. No. 82;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
DB 252 SYDSSNPV 259

RESULT 6

T13937 plexin A - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000

C/Accession: T13937

R/Windberg, M.L.; Noordermeer, J.N.; Tamagnone, L.; Comoglio, P.M.; Spriggs, M.K.; Tessie Cell 95, 903-916, 1998

A/Title: Plexin A is a neuronal semaphorin receptor that controls axon guidance.

A/Reference number: Z17621; MUID:99091049; PMID:9875845

A/Accession: T13937

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1945 <MIN>

A/Cross-references: EMBL:AF106932; NID:94056673; PID:94056674; PIDN:AAD09425.1

C/Genetics:

A:Gene: plexA

A/Cross-references: FlyBase:FBgn0025741

A/Map position: 4

C/Function:

A/Description: may function as repellents during axon guidance

C/Keywords: cell adhesion; nerve

Query Match 84.6%; Score 33; DB 2; Length 1945;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
DB 1266 SYDSSNII 1273

RESULT 7

L6H048 Ig lambda chain V-VI region (Nig-48) - human

C/Species: Homo sapiens (man)

C/Date: 28-Feb-1980 #sequence_revision 28-Feb-1980 #text_change 02-Sep-1997

```

C:Accession: A01991
R:Takahashi, N.; Takayasu, T.; Isebe, T.; Shinoda, T.; Okuyama, T.; Shimizu, A.
J. Biochem. 86, 1523-1535, 1979
A>Title: Comparative study on the structure of the light chains of human immunoglobulins
A:Reference number: A01991; MUID:80094390; PMID:118171
A:Accession: A01991
A:Molecule type: protein
A:Residues: 1-112 <TAK>
C>Note: This is the first sequenced V region of lambda chain subgroup VI
C/Comment: This is a Hence Jones protein.
A:Gene: GDB:ICLV@
C:Genetics:
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap)
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer
F:15-93/Domain: immunoglobulin homology <IMM>
F:22-91/Disulfide bonds: #stratus predicted

Query Match
Best Local Similarity      82.1%; Score 32; DB 1; Length 112;
Matches 6; Conservative    1; Pred. No. 12;
                               Mismatches 0; Indels 0; Gaps 0;

OY      1 YDSSNV 7
Db       93 YDSSLV 99
          |||||:
          |||||:

RESULT 8
B37837
Probable alpha-amylase (EC 3.2.1.1) - Clostridium acetobutylicum (Fragment)
C:Species: Clostridium acetobutylicum
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 22-Jun-1999
C:Accession: B37837
R:Gerischer, U.; Duerre, P.
J. Bacteriol. 172, 6907-6918, 1990
A>Title: Cloning, sequencing, and molecular analysis of the acetacetate decarboxylase g
A:Reference number: A37837; MUID:91072241; PMID:2254264
A:Accession: B37837
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-469 <GER>
A:Cross-references: GB:M55392; GB:M34078; NID:g144708; PIDN:AAA63759.1; PID:g144709
C:Superfamily: alpha-amylase, subtilisin type; alpha-amylase core homology
C:Keywords: glycosidase; hydrolase

Query Match
Best Local Similarity      82.1%; Score 32; DB 2; Length 469;
Matches 6; Conservative    1; Pred. No. 57;
                               Mismatches 0; Indels 0; Gaps 0;

OY      2 YDSSNV 8
Db       286 YDSSLV 292
          |||||:
          |||||:

RESULT 9
B88040
Protein F47F6.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B88040
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99065613; PMID:9851916
A>Note: see websites genome.musc1.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele
A>Note: published extra appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88040
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-527 <STO>
```

```

A:Cross-references: GB:chr_II, PIDN:AC71112.1; PID:gl707160; GSPDB:GN00020; CESP:F47F6.3
C:Genetics:
A:Gene: F47F6.3
A:Map position: 2

Query Match
Best Local Similarity 75.0%; Score 32; DB 2; Length 527;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSNNV 8
|||||:
Db 337 SYDSNNAL 344

RESULT 10
558356
pept protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C:Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 02-Feb-2001
C:Accession: S66651; S58356
R: Meyer, C.; Bierbaum, G.; Heldrich, C.; Reis, M.; Sueling, J.; Iglesias-Wind, M.I.; Kem
Bur, J. Biochem. 222, 478-489, 1995
A:Title: Nucleotide sequence of the lantibiotic Pep5 biosynthetic gene cluster and funct
A:Reference number: S66651; MUID:96035882; PMID:7556197
A:Accession: S66651
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-571 <ME2>
A:Cross-references: EMBL:Z49865; NID:g945014; PIDN:CAA90021.1; PID:g945015
C:Genetics:
A:Gene: pept
C:Superfamily: Escherichia coli ABC transporter mdla; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:348-543/Domain: ATP-binding cassette homology <ABC>
F:365-372/Region: nucleotide-binding motif A (P-loop)

Query Match
Best Local Similarity 82.1%; Score 32; DB 2; Length 571;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSNNV 8
|||||:
Db 341 SYDSNNVL 348

RESULT 11
T20762
hypothetical protein FltC1.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20762
R: Palmer, S.
submitted to the EMBL Data Library, September 1995
A:Reference number: Z19321
A:Accession: T20762
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-130 <MIT>
A:Cross-references: EMBL:Z54270; PIDN:CAA91026.1; GSPDB:GN00028; CESP:FltC1.2
A:Experimental source: clone FltC1
C:Genetics:
A:Gene: CESP:FltC1.2
A:Map position: X
A:introns: 23/3; 63/3; 89/1

Query Match
Best Local Similarity 79.5%; Score 31; DB 2; Length 130;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSNNV 7
|||||:
Db 68 SFDSNNV 74

```

RESULT 12
A30154
IGF receptor alpha chain precursor - rat
N:Alternate names: Fc-epsilon-R alpha chain precursor
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 21-Jan-2000
C/Accession: C31327; A30154; A27116; I55304
R:Lin. F.T.; Albrandt, K.; Robertson, M.W.
Proc. Natl. Acad. Sci. U.S.A. 85, 5639-5643, 1988
A>Title: cDNA heterogeneity suggests structural variants related to the high-affinity IGF
A/Reference number: A94203; MUID:88289772; PMID:2965994
A/Accession: C31327
A:Molecule type: mRNA
A:Residues: 1-245 <L1U>
A/Cross-references: GB:M21622; GB:J03811
A:Experimental source: basophilic leukemia cell line, clone R3-4
A/Accession: A31327
A:Molecule type: mRNA
A:Residues: 21-245 <L13>
A/Cross-references: GB:M21622; NID:g204109; PIDN:AAA41146.1; PID:g204110; GB:J03811
A:Experimental source: basophilic leukemia cell line
R:Shimizu, A.; Tepfer, I.; Bentley, P.N.; Berenstein, E.H.; Siraganian, R.P.; Leder, P.
Proc. Natl. Acad. Sci. U.S.A. 85, 1907-1911, 1988
A>Title: Human and rat mast cell high-affinity immunoglobulin E receptors: characterization
A/Reference number: A94191; MUID:88158102; PMID:2964640
A/Accession: A30154
A:Molecule type: mRNA
A:Residues: 1-245 <SHI>
A/Cross-references: GB:J03606; NID:g205331; PIDN:AAA41582.1; PID:g205332
R:Kinet, J.P.; Metzger, H.; Hakim, J.; Kochan, J.
Biochemistry 26, 4605-4610, 1987
A>Title: A cDNA presumptively coding for the alpha subunit of the receptor with high aff
A/Reference number: A27116; MUID:88024987; PMID:2959318
A/Accession: A27116
A:Molecule type: mRNA
A:Residues: 1, 'G', '3-236, 'N', 238-244, 'RLKENS', <KIN>
R:Tepfer, I.; Shimizu, A.; Leder, P.
J. Biol. Chem. 264, 5912-5915, 1989
A>Title: The gene for the rat mast cell high affinity IGE receptor alpha chain. Structure
A/Reference number: I55304; MUID:89174653; PMID:2522441
A/Accession: I55304
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A/Cross-references: GB:M25334; NID:g341335; PIDN:AAA74562.1; PID:g556391
C/Superfamily: Fc gamma receptor II; Immunoglobulin homology
C/Keywords: immunoglobulin receptor; transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-245/Product: IGE receptor alpha chain #status predicted <MAT>
F:42-93/Domain: immunoglobulin homology <IMV>

Query Match 79.5%; Score 31; DB 2; Length 245;
Best Local Similarity 71.4%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSNNV 7
DB 153 SYDSNNI 159

RESULT 13
G84781
hypothetical protein At2g35540 (imported) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: G84781
R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Neues, D.; Nleman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Genus 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: G84781
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-249 <STO>
A/Cross-references: GB:AE002093; NID:g4581152; PIDN:AA024636.1; GSPDB:GN00139
C/Genetics:
A:Gene: At2g35540
A/Map position: 2

Query Match 79.5%; Score 31; DB 2; Length 249;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSNNV 7
DB 187 SYDSNNI 193

RESULT 14
T15520
hypothetical protein C15H9.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C/Accession: T15520
R:Bentley, D.
submitted to the EMBL Data Library, April 1996
A>Description: The sequence of C. elegans cosmid C15H9.
A/Reference number: Z18364
A/Accession: T15520
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-261 <BEN>
A/Cross-references: EMBL:U56965; NID:g1945492; PID:g1293841; PIDN:AA052662.1; GSPDB:GN00
A:Experimental source: strain Bristol N2; clone C15H9
C/Genetics:
A:Gene: CESP:C15H9.2
A/Map position: X
A:Introns: 16/3; 40/2; 78/3; 109/2; 137/3; 192/1; 244/3

Query Match 79.5%; Score 31; DB 2; Length 261;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSNNV 8
DB 122 FDSNNV 128

RESULT 15
JC4584
insulin-like growth factor binding protein-5 precursor - pig
C/Species: Sus scrofa domestica (domestic pig)
C>Date: 10-Apr-1996 #sequence_revision 24-May-1996 #text_change 15-Jul-2002
C/Accession: JC4584; G23734
R:White, M.E.; Diao, R.; Hathaway, M.R.; Mickelson, J.; Dayton, W.R.
Biochem. Biophys. Res. Commun. 218, 248-253, 1996
A>Title: Molecular cloning and sequence analysis of the porcine insulin-like growth fact
A/Reference number: JC4584; MUID:96136309; PMID:8573141
A/Accession: JC4584
A:Molecule type: mRNA
A:Residues: 1-271 <WHI>
A/Cross-references: GB:U41340; NID:g1173906; PIDN:AA07859.1; PID:g1173907
A:Experimental source: skeletal muscle
R:Shimasaki, S.; Gao, L.; Shimomura, M.; Ling, N.
Mol. Endocrinol. 5, 938-948, 1991
A>Title: Isolation and molecular cloning of insulin-like growth factor-binding protein-6
A/Reference number: A23734; MUID:92049376; PMID:1719383
A/Accession: G23734
A:Molecule type: protein
A:Residues: 20-25, 'X', 27-28, 'X', 30-36, 'X', 38-39 <SHI>
A/Comment: This protein has essential roles in the regulation and coordination of insulin
lays a role during myoblast proliferation and differentiation, and is important in the g
C/Superfamily: insulin-like growth factor binding protein 1; thyroglobulin type I repeat

C/Keywords: differentiation; skeletal muscle
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-271/Product: insulin-like growth factor binding protein-5 #status experimental <MAT
 F:191-262/Domain: thyroglobulin type I repeat homology <THYI>

Query Match 79.5%; Score 31; DB 2; Length 271;
 Best Local Similarity 85.7%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDSSNV 7
 |.|||||
 Db 264 SPDSSNV 270

Search completed: January 12, 2004, 07:03:10
 Job time : 10.25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:49 ; Search time 4.375 Seconds
(without alignments)
85.992 Million cell updates/sec

Title: US-09-829-495-66

Perfect score: 39
Sequence: 1 SYDSNNV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	84.6	617	1 VAA1_MOUSE	P50516 mus musculu
2	32	82.1	112	1 LV6B_HUMAN	P01722 homo sapien
3	32	82.1	460	1 COA1_BPIF1	O80297 bacterioph
4	32	82.1	760	1 AMY_CLOAB	P23671 clostridium
5	31	79.5	245	1 FCE1_RAT	P12371 rattus norv
6	31	79.5	271	1 IBB5_PIG	O28965 sus scrofa
7	31	79.5	272	1 SER3_DROME	P17207 drosophila
8	31	79.5	514	1 FOLE_CANAL	O97893 candida alb
9	31	79.5	548	1 PRZN_RENSA	P55111 renbacteri
10	31	79.5	963	1 CHS2_YEAST	P14180 saccharomyc
11	31	79.5	964	1 CTA2_BACCI	P70873 bacillus ci
12	30	76.9	332	1 REGA_CLOSA	O45831 clostridium
13	30	76.9	429	1 YN8B_YEAST	P53769 saccharomyc
14	30	76.9	462	1 MB12_ARATH	O98400 arabidopsis
15	30	76.9	624	1 VP87_NPYOP	P17930 oryza pseu
16	30	76.9	630	1 V109_POMPV	P36317 fowipox vir
17	30	76.9	754	1 MBS3_CAREL	O10665 caenorhabdi
18	30	76.9	769	1 ITB2_HUMAN	P05107 homo sapien
19	30	76.9	798	1 ITB7_HUMAN	P26107 homo sapien
20	30	76.9	803	1 GYRB_BUCAI	P57126 buchera ap
21	30	76.9	806	1 ITB7_MOUSE	P26011 mus musculu
22	30	76.9	816	1 COLA_VIBPA	O56666 vibrio para
23	30	76.9	1483	1 YCDA_DROME	O98941 drosophila
24	29	74.4	284	1 YHDO_ECOLI	P28638 escherichia
25	29	74.4	483	1 YCUD_YEAST	P11230 saccharomyc
26	29	74.4	501	1 ACHB_HUMAN	P11230 homo sapien
27	29	74.4	501	1 ACHB_MOUSE	P09690 mus musculu
28	29	74.4	501	1 ACHB_RAT	P25109 rattus norv
29	29	74.4	505	1 ACHB_BOVIN	P04378 bos taurus
30	29	74.4	528	1 FOLE_NEBCR	O13442 neotropora
31	29	74.4	758	1 YP58_YEAST	O99259 saccharomyc
32	29	74.4	762	1 SIAP_ACEKI	P22258 acetogenium
33	29	74.4	771	1 ITB2_MOUSE	P11835 mus musculu

34	29	74.4	775	1 CDR2_SCHPO	P87050 schizosacch
35	29	74.4	784	1 LON_RICPR	O92492 rickettsia
36	29	74.4	803	1 GYRB_ECOLI	P06982 escherichia
37	29	74.4	803	1 GYRB_SALTY	O60008 salmonella
38	29	74.4	1916	1 RIF1_YEAST	P29539 saccharomyc
39	28	71.8	173	1 LEC2_MEGRO	P17346 megabalanus
40	28	71.8	209	1 RH01_YEAST	P06780 saccharomyc
41	28	71.8	221	1 EPD2_ONCMY	P28771 oncorhynch
42	28	71.8	230	1 Y920_HELPY	O92871 helicobacte
43	28	71.8	230	1 Y920_HELPY	O25578 helicobacte
44	28	71.8	261	1 ARSA_PHAUV	O42460 phaseolus v
45	28	71.8	261	1 ARSB_PHAUV	O41116 phaseolus v

ALIGNMENTS

RESULT 1
ID VAA1_MOUSE STANDARD, PRT, 617 AA.
AC P50516;
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Vacuolar ATP synthase catalytic subunit A, ubiquitous isoform
DE (EC 3.6.3.14) (V-ATPase A subunit 1) (Vacuolar proton pump alpha
DE subunit 1) (V-ATPase 69 kDa subunit 1).
CN ATP6V1A OR ATP6V1A1 OR ATP6A1 OR ATP6A2.
OS Mus musculus (Mouse).
OC Mammalia; Chordata; Crustacea; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96362668; PubMed=8741845;
RA Laitala T., Howell M.L., Dean G.E., Vaananen H.K.;
RT "Resorption-cycle-dependent polarization of mRNAs for different
RT subunits of V-ATPase in bone-resorbing osteoclasts.";
RL Mol. Biol. Cell 7:129-142(1996).
CC -1- FUNCTION: CATALYTIC SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF
CC VACUOLAR ATPASE. V-ATPASE VACUOLAR ATPASE IS RESPONSIBLE FOR
CC ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC
CC CELLS.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U13837; AAC52410.1; -
CC MGD: MG1:1201780; Atb6v1a1.
DR InterPro: IPR000793; ATPase_a/bc.
DR InterPro: IPR000194; ATPase_a/bc.
DR InterPro: IPR004100; ATPase_a/bn.
DR InterPro: IPR005725; V_ATPase_V1_A.
DR Pfam: PF00006; ATP-synt_ab; 1.
DR Pfam: PF00306; ATP-synt_ab C; 1.
DR Pfam: PF02874; ATP-synt_ab N; 1.
DR TRIGRAMS: TRIGRAM1042; V-ATPase_V1_A; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; 1.
KW ATP synthase; Hydrogen ion transport; Hydrolyase; ATP-binding;
KW Multigene family.
FT NP_BIND 250 257 ATP (POTENTIAL).

SQ SEQUENCE 617 AA; 68268 MW; 778D2843070200CF CRC64;
 Query Match 84.6%; Score 33; DB 1; Length 617;
 Best Local Similarity 75.0%; Pred. No. 18;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
 DB 184 NYDASNV 191

RESULT 2
 LVB6 HUMAN STANDARD; PRT; 112 AA.
 ID LV6B_HUMAN
 AC P01722;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE IG lambda chain V-VI region NIG-48.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=80094390; PubMed=118171;
 RA Takahashi N., Takayasu T., Isebe T., Shinoda T., Okuyama T.,
 RA Shimizu A.;
 RT "Comparative study on the structure of the light chains of human
 RT immunoglobulin. II. Assignment of a new subgroup."
 RL J. Biochem. 86:1523-1535(1979).
 CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 PR, A01991; LSHU8.
 DR HSSP; P01703; 7FAB.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_V.
 DR Pfam; PF00047; IG_1.
 DR SMART; SM00406; IGV_1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KM Immunoglobulin V region; Bence-Jones protein.
 FT DOMAIN 1 107 IG-LIKE.
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 12152 MW; CFB307BC527A384 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 112;
 Best Local Similarity 85.7%; Pred. No. 42;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
 DB 93 SYDSSNV 99

RESULT 3
 COA6 BP1F1 STANDARD; PRT; 460 AA.
 ID COA6_BP1F1
 AC O80297;
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Coat protein A precursor (Attachment protein).
 OS Ili OR 3.
 OS Bacteriophage If1.
 OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
 ON NCB1_TaxID=10866;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hill D.F., Hughes G., McNaughton J.C., Stockwell P.A., Petersen G.B.;

RT "DNA sequence of the filamentous coliphage If1."
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: COAT PROTEIN A IS NECESSARY FOR ADSORPTION OF THE VIRION
 CC ONTO THE F-PILUS OF THE HOST CELL (BY SIMILARITY).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U02303; AAC82155.1; -.
 DR EMBL; P03661; 1FGB.
 DR Pfam; PF05357; Phage.Coat A; 1.
 KM Phage recognition; Coat protein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 460 COAT PROTEIN A.
 SQ SEQUENCE 460 AA; 48790 MW; 9AD651968C715AB2 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 460;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
 DB 229 SYDTSNV 235

RESULT 4
 AMY CLOAB STANDARD; PRT; 760 AA.
 ID AMY_CLOAB
 AC P23671; Q9S429;
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase).
 GN AMYA OR AMYP OR CAP0168.
 OS Clostridium acetobutylicum.
 OG Plaemid PSOLI.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 ON NCB1_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RT Schaefer S., Duerr P.;
 RT "Nucleotide sequence analysis and ECF sigma factor-dependent
 RT expression of an alpha-amylase gene from Clostridium acetobutylicum
 RT DSM 792."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RA Sabache F., Cornillot E., Croux C., Soucaille P.;
 RT "Molecular characterization of amyP, a psoli located gene coding the
 RT major alpha-amylase of Clostridium acetobutylicum ATCC824, and its use
 RT as a reporter system for strain degeneration."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R.L., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabache F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum."
 RL J. Bacteriol. 183:4823-4838(2001).
 RN [4]

RP SEQUENCE OF 292-760 FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=91072241; PubMed=2254264;
 RA Gerischer U., Duerre P.;
 RT "Cloning, sequencing, and molecular analysis of the acetate
 decarboxylase gene region from Clostridium acetobutylicum";
 RL J. Bacteriol. 172:6907-6918(1990).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 linkages in oligosaccharides and polysaccharides.
 CC SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M55392; AAA63759.2; -
 DR EMBL: AF164199; AAD47072.1; -
 DR EMBL: AE001438; AAK76913.1; -
 DR PIR: B37837; B37837.
 DR HSSP: P00691; 1BAG.
 DR InterPro: IPR006048; Alpha_ami1_C.
 DR InterPro: IPR006047; Alpha_ami1_cat.
 DR InterPro: IPR005085; CBM_25.
 DR InterPro: IPR006046; Glyco_hydro_13.
 DR Pfam: PF00128; alpha-amy1ase_1.
 DR Pfam: PF02806; alpha-amy1ase_C_1.
 DR Pfam: PF03423; CBM_25; 1.
 DR PRINTS: PRO0110; ALPHAMYLASE.
 DR SMART: SM00632; Amy_C_1.
 KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Signal; Plasmid;
 KM Complete proteome.
 FT SIGNAL 1 34 POTENTIAL.
 FT CHAIN 35 760 ALPHA-AMYLASE.
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT ACT_SITE 226 226 BY SIMILARITY.
 FT ACT_SITE 262 262 BY SIMILARITY.
 FT ACT_SITE 321 321 BY SIMILARITY.
 FT CONFLICT 213 213 D -> H (IN REF. 1).
 FT CONFLICT 222 222 D -> H (IN REF. 1).
 FT CONFLICT 571 571 A -> G (IN REF. 1).
 FT SEQUENCE 760 AA; 83045 MW; 10E146F40252FED0 CRC64;
 SO
 Query Match 82.1%; Score 32; DB 1; Length 760;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RA Kinet J.-P., Metzger H., Hakim J., Kochan J.;
 RT "A cDNA presumably coding for the alpha subunit of the receptor
 RT with high affinity for immunoglobulin E.";
 RL Biochemistry 26:4605-4610(1987).
 RN [2]
 RN REVISIONS.
 RA Kochan J.;
 RL Submitted (MAR-1988) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=rat cells;
 RX MEDLINE=88158102; PubMed=2964640;
 RA Shimizu A., Teppler I., Bentley P.N., Siraganian R.P.,
 RA Leder P.;
 RT "Human and rat mast cell high-affinity immunoglobulin E receptors:
 RT characterization of putative alpha-chain gene products.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988).
 RN [4]
 RP SEQUENCE OF 21-245 FROM N.A.
 RX MEDLINE=88289772; PubMed=2969594;
 RA Liu F.-T., Albrandt K., Robertson M.W.;
 RT "cDNA heterogeneity suggests structural variants related to the high-
 RT affinity IgE receptor";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:5639-5643(1988).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
 CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
 CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IgE LEADS TO CELL
 CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
 CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
 CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
 CC DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M17153; AAA42045.1; -
 DR EMBL: J03606; AAA41582.1; -
 DR EMBL: M21622; AAA41146.1; -
 DR PIR: C31327; A30154.
 DR HSSP: P12319; 1ALS.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_2.
 DR SMART: SM00409; Ig_2.
 DR PROSITE: PS50835; Ig_LIKE; 1.
 KW IgE-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KM Immunoglobulin domain; Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 245
 FT DOMAIN 24 245 HIGH AFFINITY IMMUNOGLOBULIN EPSILON
 FT DOMAIN 205 223 RECEPTOR ALPHA-SUBUNIT.
 FT DOMAIN 224 245 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 28 103 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 113 181 IG-LIKE 1.
 FT DISULFID 49 91 IG-LIKE 2.
 FT DISULFID 130 174 BY SIMILARITY.
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 245 AA; 27793 MW; A06E7DD363B72197 CRC64;
 Query Match 79.5%; Score 31; DB 1; Length 245;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
 153 SYDSSNMI 159

RESULT 6
 ID IDB5_PIG STANDARD; PRT; 271 AA.
 AC 0289B5;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein 5 precursor (IGFBP-5)
 DE (IGFBP-5) (IGF-binding protein 5).
 GN IGFBP5.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skletal muscle;
 RX MEDLINE=96136309; PubMed=8573141;
 RA White M.E., Diao R., Hathaway M.R., Mickelson J., Dayton W.R.;
 RT "Molecular cloning and sequence analysis of the porcine insulin-like
 RT growth factor binding protein-5 complementary deoxyribonucleic
 RT acid";
 RL Biochem. Biophys. Res. Commun. 218:248-253(1996).
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 thyroglobulin type-I domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U41340; AAA87859.1; -.
 DR PIR; J04584; J04584.
 DR HSP; P24593; IBOE.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR00716; Thyroglobulin_1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00066; thyroglobulin_1; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00211; TY; 1.
 DR PROSITE; PS00422; IGF_BINDING; 1.
 DR PROSITE; PS00484; THYROGLOBULIN_1; 1.
 DR Growth factor binding; signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 271 INSULIN-LIKE GROWTH FACTOR BINDING
 FT DOMAIN 214 262 THYROGLOBULIN TYPE I.
 FT SEQUENCE 271 AA; 30323 MW; 95C254E95DC4EDBD CRC64;

Query Match 79.5%; Score 31; DB 1; Length 271;
 Best Local Similarity 85.7%; Pred. No. 19;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7

Db 264 SPDSSNV 270
 264 SPDSSNV 270

RESULT 7
 ID SER3_DROME STANDARD; PRT; 272 AA.
 AC P17207; Q9VAD8;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine protease 3 precursor (EC 3.4.21.-).
 GN SER99DC OR SER3 OR CG17951.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
 RA Abri'l J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballow R.M., Basu A., Baxendale U., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsch C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jaitai M., Kalish F., Karpem G.H., Ke Z., Kenton J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mikhlov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclik J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE OF 212-272 FROM N.A.
 RX MEDLINE=89219063; PubMed=2469005;
 RA Yun Y., Davis R.L.;
 RT "Levels of RNA from a family of putative serine protease genes are
 RT reduced in Drosophila melanogaster dunce mutants and are regulated by
 RT cyclic AMP";
 RL Mol. Cell. Biol. 9:692-700(1989).
 CC -1- FUNCTION: ITS MAJOR FUNCTION MAY BE TO AID IN DIGESTION.
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THE LARVAL GUT.
 CC -1- DEVELOPMENTAL STAGE: BEGAN TO APPEAR AT LATE EMBRYO STAGE AND
 CC CONTINUED TO INCREASE IN ABUNDANCE THROUGHOUT THE LARVAL STAGE.
 CC THEY ARE NOT PRESENT IN PUPAE BUT REAPPEARED IN THE ADULT.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB003771; AAF56973.1; -
 CC EMBL: M24380; AAB02551.1; -
 CC PIR: PS0049; PS0049.
 CC HSSP: P00734; 1B7X.
 CC MEROPS: S01.UPA; -
 CC PDBase: Pfam0003358; Ser99DC.
 CC InterPro: IPR001314; Chymotrypsin.
 CC InterPro: IPR001254; Ser_protease_Try.
 CC Pfam: PF00089; trypsin_1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00020; TRYP_SPC; 1.
 CC PROSITE: PS00134; TRYP_SIN_DOM; 1.
 CC PROSITE: PS00135; TRYP_SIN_SER; 1.
 CC K0 Hydrolyase; Serine protease; Signal; Zymogen; Multigene family.
 CC FT SIGNAL 1 17
 CC FT PROPEP 18 40
 CC FT CHAIN 41 272
 CC FT ACT_SITE 84 84
 CC FT ACT_SITE 127 127
 CC FT ACT_SITE 222 222
 CC FT DISULFID 69 85
 CC FT DISULFID 193 208
 CC FT DISULFID 218 246
 CC FT CONFLICT 212 212
 CC SQ SEQUENCE 272 AA; 29620 MW; A61B9DE3AFC93AF CRC64;
 CC
 CC Query Match 79.5%; Score 31; DB 1; Length 272;
 CC Best Local Similarity 85.7%; Pred. No. 19;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 YDSSNVV 8
 CC Db 171 YDGSNVV 177
 CC
 CC RESULT 8
 CC ID FOLE CANAL STANDARD; PRT; 514 AA.
 CC AC 09Y853;
 CC DT 15-SEP-2003 (Rel. 42, Created)
 CC DT 15-SEP-2003 (Rel. 42, Last sequence update)
 CC DE FOLYPOLYGLUTAMATE SYNTHASE (EC 6.3.2.17) (FOLYPOLY-GAMMA-GLUTAMATE
 CC synthetase) (FPGS).
 CC GN MET7.
 CC OS Candida albicans (Yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 CC OX NCBI_TaxID=5476;
 CC OX [1]
 CC RP SEQUENCE FROM N.A.
 CC RA Subramanian M., Li C., DeSouza L., Bogner A.L.;
 CC RT "Sequencing of a Candida albicans genomic library plasmid which
 CC encodes the complete open reading frame of folypolyglutamate
 CC synthetase".
 CC RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC CC -1- FUNCTION: Conversion of folates to polyglutamate derivatives (By
 CC similarity).
 CC CC -1- CATALYTIC ACTIVITY: ATP + (tetrahydrofolyl)-[Glu] (N) + L-glutamate
 CC = ADP + phosphate + (tetrahydrofolyl)-[Glu] (N+1).
 CC CC -1- SIMILARITY: Belongs to the folypolyglutamate synthase family.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AF156928; AAD40312.1; -
 CC HSSP: P15925; 1FGS.
 CC InterPro: IPR001645; Fpolylgl_synthase.
 CC TIGRFA: TIGR01499; foic_1.
 CC PROSITE: PS01011; FOLYPOLYGLU_SYNT_1; FALSE_NEG.
 CC PROSITE: PS01012; FOLYPOLYGLU_SYNT_2; 1.
 CC KM Lysase; One-carbon metabolism; ATP-binding.
 CC NP BIND 78 84
 CC FT NP BIND 78 84
 CC SQ SEQUENCE 514 AA; 57907 MW; EABF3C8CC79A12A CRC64;
 CC
 CC Query Match 79.5%; Score 31; DB 1; Length 514;
 CC Best Local Similarity 50.0%; Pred. No. 39;
 CC Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 SYDSSNVV 8
 CC Db 192 TYDSTNII 199
 CC
 CC RESULT 9
 CC ID PRZN RENSA STANDARD; PRT; 548 AA.
 CC AC .P5511;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Zinc metalloproteinase precursor (EC 3.4.24.-) (Hemolysin).
 CC GN HLY.
 CC OS Renibacterium salmoninarum.
 CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC OC Micrococcales; Micrococcaceae; Renibacterium.
 CC OX NCBI_TaxID=1646;
 CC OX [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=MT44;
 CC MEDLINE=95400487; PubMed=7545509;
 CC RA Grayson T.H., Evenden A.J., Gilpin M.L., Martin K.L., Munn C.B.;
 CC RT "A gene from Renibacterium salmoninarum encoding a product which
 CC shows homology to bacterial zinc-metalloproteases".
 CC RL Microbiology 141:1331-1341(1995).
 CC RN [2]
 CC RP REVISIONS TO 149-170.
 CC RA Grayson T.H.;
 CC RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC CC -1- FUNCTION: ZINC METALLOPROTEASE WITH HEMOLYTIC PROPERTIES.
 CC CC HEMOLYTIC ACTIVITY IS OBSERVED FROM 6 TO 37 DEGREES CELSIUS FOR
 CC MAMMALIAN ERYTHROCYTES.
 CC CC -1- COPACITOR: Binds 1 zinc ion and 4 calcium ions per subunit (By
 CC similarity).
 CC CC -1- SUBCELLULAR LOCATION: NOT SECRETED, BUT PROBABLY REMAINS ATTACHED
 CC OR ASSOCIATED WITH THE CELL WALL.
 CC CC -1- INDUCTION: EXPRESSION OF THE HEMOLYSIN IS MODULATED BY THE
 CC AVAILABILITY OF IRON.
 CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X76499; CAA54032.2; -
 CC HSSP: P00800; 1HYT.

DR MEROPS; M04.UPW; -
 DR InterPro; IPR001570; Peptidase M4.
 DR InterPro; IPR006025; Zn MTpeptidase.
 DR Pfam; PF01447; Peptidase_M4.1.
 DR Pfam; PF02868; Peptidase_M4.C1.
 DR PRINTS; PR00730; THERMOLYSIN.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 KM Hydroxylase; Metalloprotease; Zinc; Calcium; Zymogen; Signal.
 FT SIGNAL 1 28
 FT PROPEP 29 ?
 FT CHAIN 29 548
 FT METAL 362 362
 FT ACT SITE 363 363
 FT METAL 366 366
 FT METAL 386 386
 FT ACT SITE 459 459
 SQ SEQUENCE 548 AA; 56378 MW; 7AD090DD7425B90A CRC64;
 Query Match 79.5%; Score 31; DB 1; Length 548;
 Best Local Similarity 87.5%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SYSDSNV 8
 Db 35 SLDSNV 42
 RESULT 10
 CHS2_YEAST STANDARD; PRT; 963 AA.
 ID CHS2_YEAST
 AC P14180;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl transferase 2).
 GN CHS2 OR YBR038W OR YBR0407.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9014317; PubMed=2533436;
 RA Silverman S.J.;
 RT "Similar and different domains of chitin synthases 1 and 2 of S. cerevisiae: two isozymes with distinct functions.";
 RL Yeast 5:459-467(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPRAIN=5288C;
 RA Andre B., Cziepluch C., Hein C., Jauniaux J.C., Ureastazu A., Vissers S.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP MUTAGENESIS.
 RX MEDLINE=9529398; PubMed=775457;
 RA Nagashashi S., Sudon M., Ono N., Sawada R., Yamaguchi E., Uchida Y., Mo T., Takagi M., Aisawa M., Yamada-Okabe H.;
 RT "Characterization of chitin synthase 2 of Saccharomyces cerevisiae. Implication of two highly conserved domains as possible catalytic sites.";
 RT Biol. Chem. 270:13961-13967(1995).
 CC -1- FUNCTION: ESSENTIAL FOR SEPTUM FORMATION AND CELL DIVISION. CHS2 IS REQUIRED FOR MAINTAINING NORMAL CELL MORPHOLOGY.
 CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + ((1,4)-(N-acetyl-beta-D-glucosaminyl)) (N) = UDP + ((1,4)-(N-acetyl-beta-D-glucosaminyl)) (N+1).
 CC -1- ENZYME REGULATION: REQUIRES PROTEOLYTIC ACTIVATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M23865; AAA34493.1; -
 DR EMBL; Z35907; CAA84980.1; -
 DR PIR; S45167; S45167.
 DR SGD; S0000242; CHS2.
 DR GO; GO:0001442; Contractile ring (sensu Saccharomycetes); IDA.
 DR GO; GO:0004100; F-actin synthase activity; IDA.
 DR GO; GO:0016288; P-cytokinesis; IMP.
 DR GO; GO:0006970; P-response to osmotic stress; IMP.
 DR InterPro; IPR004834; Chitin synth.
 DR InterPro; IPR001173; Glyco Trans_2.
 DR Pfam; PF01644; Chitin synth. 1.
 DR Prodom; PD002998; Chitin synth; 1.
 KM Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 KM Multigene family.
 FT DOMAIN 1 422
 FT TRANSMEM 423 443
 FT DOMAIN 444 643
 FT TRANSMEM 644 664
 FT DOMAIN 665 677
 FT TRANSMEM 678 698
 FT DOMAIN 699 711
 FT TRANSMEM 712 732
 FT DOMAIN 733 743
 FT TRANSMEM 744 764
 FT DOMAIN 765 775
 FT TRANSMEM 776 796
 FT DOMAIN 797 875
 FT TRANSMEM 876 896
 FT DOMAIN 897 905
 FT TRANSMEM 906 926
 FT DOMAIN 927 963
 FT TRANSMEM 197 197
 FT TRANSMEM 312 312
 FT TRANSMEM 355 355
 FT TRANSMEM 393 393
 FT TRANSMEM 441 441
 FT TRANSMEM 441 441
 FT TRANSMEM 447 447
 FT TRANSMEM 490 490
 FT TRANSMEM 492 492
 FT TRANSMEM 493 493
 FT TRANSMEM 494 494
 FT TRANSMEM 497 497
 FT TRANSMEM 502 502
 FT TRANSMEM 505 505
 FT TRANSMEM 506 506
 FT TRANSMEM 508 508
 FT TRANSMEM 509 509
 FT TRANSMEM 514 514
 FT TRANSMEM 515 515
 FT TRANSMEM 516 516
 FT TRANSMEM 521 521
 FT TRANSMEM 522 522
 FT TRANSMEM 537 537
 FT TRANSMEM 550 550
 FT TRANSMEM 556 556
 FT TRANSMEM 559 559
 FT TRANSMEM 561 561
 FT TRANSMEM 561 561
 FT TRANSMEM 562 562
 FT TRANSMEM 562 562
 FT TRANSMEM 563 563
 FT TRANSMEM 563 563
 FT TRANSMEM 565 565
 FT TRANSMEM 565 565
 EXTRACELLULAR (POTENTIAL).
 1 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 2 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 3 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 4 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 5 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 6 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 7 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 8 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 N->A: 30% DECREASE OF ACTIVITY.
 S->A: 20% INCREASE OF ACTIVITY.
 D->A: 10% DECREASE OF ACTIVITY.
 D->A: 5% DECREASE OF ACTIVITY.
 D->A: LOSS OF ACTIVITY.
 D->E: LOSS OF ACTIVITY.
 N->A: 80% DECREASE OF ACTIVITY.
 O->A: 70% DECREASE OF ACTIVITY.
 F->A: 80% DECREASE OF ACTIVITY.
 E->A: 90% DECREASE OF ACTIVITY.
 Y->A: 95% DECREASE OF ACTIVITY.
 S->A: 95% DECREASE OF ACTIVITY.
 K->A: 90% DECREASE OF ACTIVITY.
 E->A: 80% DECREASE OF ACTIVITY.
 F->A: 20% INCREASE OF ACTIVITY.
 S->A: 60% DECREASE OF ACTIVITY.
 L->A: 80% DECREASE OF ACTIVITY.
 G->A: 90% DECREASE OF ACTIVITY.
 Y->A: 95% DECREASE OF ACTIVITY.
 R->A: 60% DECREASE OF ACTIVITY.
 R->A: NO CHANGE IN ACTIVITY.
 H->A: 85% DECREASE OF ACTIVITY.
 N->A: 95% DECREASE OF ACTIVITY.
 L->A: 95% DECREASE OF ACTIVITY.
 E->A: LOSS OF ACTIVITY.
 E->D: 65% DECREASE OF ACTIVITY.
 D->Q: LOSS OF ACTIVITY.
 E->A: LOSS OF ACTIVITY.
 D->E: LOSS OF ACTIVITY.
 D->N: LOSS OF ACTIVITY.
 R->A: LOSS OF ACTIVITY.
 L->A: 95% DECREASE OF ACTIVITY.


```

FT MUTAGEN 589 589 T->A: 70% DECREASE OF ACTIVITY.
FT MUTAGEN 592 592 P->A: 70% DECREASE OF ACTIVITY.
FT MUTAGEN 601 601 O->A: LOSS OF ACTIVITY.
FT MUTAGEN 601 601 O->N: LOSS OF ACTIVITY.
FT MUTAGEN 602 602 R->A: 95% DECREASE OF ACTIVITY.
FT MUTAGEN 602 602 R->K: 95% DECREASE OF ACTIVITY.
FT MUTAGEN 603 603 R->A: LOSS OF ACTIVITY.
FT MUTAGEN 603 603 R->K: 57% DECREASE OF ACTIVITY.
FT MUTAGEN 602 603 RR->KK: LOSS OF ACTIVITY.
FT MUTAGEN 604 604 R->A: LOSS OF ACTIVITY.
FT MUTAGEN 604 604 R->K: LOSS OF ACTIVITY.
FT MUTAGEN 605 605 W->A: LOSS OF ACTIVITY.
FT MUTAGEN 605 605 W->Y: LOSS OF ACTIVITY.
FT MUTAGEN 607 607 N->A: 95% DECREASE OF ACTIVITY.
SQ SEQUENCE 963 AA; 109881 MW; 4D44A287C0B65B5B CRC64;

Query Match 79.5%; Score 31; DB 1; Length 963;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SYDSSNV 7
Db 47 SYDSSNV 53

RESULT 11
CTA2_BACCI STANDARD; PRT; 964 AA.
AC P70873;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Cycloisomaltooligosaccharide glucanotransferase precursor (EC 2.4.1.-)
DE (CITASE).
GN CIT.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U-155;
RA Oguma T., Kurokawa T., Tobe K., Kitao S., Kobayashi M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PRODUCES CYCLOISOMALTOOLIGOSACCHARIDE FROM DEXTRAN.
CC -1- SIMILARITY: BELONGS TO FAMILY 66 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D88360; BAA13595.1; -
CC DR InterPro: IPR005084; CBM_6.
CC DR Pfam; PF03422; CBM_6; 2_6.
CC TRANSFERASE; Glycosyltransferase; Signal.
CC FT SIGNAL 1 30 POTENTIAL.
CC FT CHAIN 31 964 GLUCANOTRANSFERASE.
CC FT SEQUENCE 964 AA; 107208 MW; 8849CC5E2DE3A68 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 964;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYDSSN 6
Db 521 SYDSSN 526

RESULT 12

```

```

REGA_CLOSA STANDARD; PRT; 332 AA.
ID REGA_CLOSA
AC Q45831;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription regulatory protein regA.
DE REGA OR REPA.
GN Clostridium saccharobutylicum.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_Taxid=169679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P262;
RX MEDLINE=95291448; PubMed=7539689;
RA Davison S.P., Santangelo J.D., Reid S.J., Woods D.R.;
RT "A Clostridium acetobutylicum regulator gene (regA) affecting amylose
RT production in Bacillus subtilis."
RL Microbiology 141:989-996 (1995).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF AMYLASE PRODUCTION.
CC -1- SIMILARITY: BELONGS TO THE LACI FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -1- CAUTION: Was originally thought to originate from
CC C.acetobutylicum.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14685; AAA23275.1; -
CC DR HSP; P15039; IPRU.
CC DR InterPro: IPR000843; HTH_Laci.
CC DR InterPro: IPR001761; PeriplaBP/Laci.
CC DR Pfam; PF00356; laci; 1.
CC DR Pfam; PF00532; Peripla_BP_like; 1.
CC DR PRINTS; PRO0036; HTHLACI.
CC DR SMART; SM00354; HTH_LACI; 1.
CC DR PROSITE; PS00356; HTH_LACI_FAMILY; 1.
CC TRANSIT: Transcription regulation; DNA-binding.
CC FT DNA_BIND 5 24 H-T-H MOTIF (POTENTIAL).
CC FT SEQUENCE 332 AA; 37260 MW; 27CB52FF9BBE007 CRC64;

Query Match 76.9%; Score 30; DB 1; Length 332;
Best Local Similarity 62.5%; Pred. No. 40;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYDSSNV 8
Db 165 SYDSTNL 172

RESULT 13
YN8H_YEAST STANDARD; PRT; 429 AA.
ID YN8H_YEAST
AC P53729;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 48.1 kDa protein in SEC12-SK2 intergenic region.
DE YNR029C OR N3560.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

```

CC -1- SIMILARITY: TO P.DENTRIFICANS COBW.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z11644; CAA96309.1; --
 CC PIR: S63360; S63360.
 CC SGI: S0005312; YNR029C.
 CC DR InterPro: IPR003495; COBW.
 CC DR Pfam: PF02492; COBW; 1.
 CC DR Hypothetical protein: ATP-binding.
 CC KW NP_BIND 76 85 ATP (POTENTIAL).
 CC SEQUENCE 429 AA; 48121 MW; 58E033E685EF764 CRC64;
 CC
 CC Query Match 76.9%; Score 30; DB 1; Length 429;
 CC Best Local Similarity 62.5%; Pred. No. 53;
 CC Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 1 SYDSNNV 8
 CC Db 36 SYDGNIV 43
 CC
 CC RESULT 14
 CC MB12 ARATH STANDARD; PRT; 462 AA.
 CC ID 09SAV0;
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Myosinase binding protein-like At1g52040.
 CC GN AT1G52040 OR F5F19.10.
 CC OS Arabidopsis thaliana (Mouse-ear cress).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 CC OX NCBI_TaxID=3702;
 CC OX [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=cv. Columbia;
 CC RX MEDLINE=21016719; PubMed=11130712;
 CC RA Theologis A., Ecker U.R., Palm C.J., Federspiel N.A., Kaul S.,
 CC RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 CC RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 CC RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
 CC RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 CC RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 CC RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,
 CC RA Kim C.J., Koo H.L., Kremetsekaia I., Kurtz D.B., Kwan A., Lam B.,
 CC RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 CC RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 CC RA Millefischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 CC RA Pai G., Peterson U., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 CC RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 CC RA Sun H., Tallon L.J., Tambunga G., Toriumi M.C., Town C.D.,
 CC RA Uetback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 CC RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.M.;
 CC RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 CC thaliana."
 CC RT Nature 408:816-820(2000).
 CC -1- SIMILARITY: BELONGS TO THE JACALIN LECTIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AC006216; AAD12678.1; --
 CC DR PIR: B96560; B96560.
 CC DR HSSP: P18670; IJAC.
 CC DR InterPro: IPR001229; Jacalin_lectin.
 CC DR Pfam: PF01419; Jacalin; 3.
 CC KW Lectin; Repeat; Multigene family.
 CC SEQUENCE 462 AA; 50167 MW; 0A6DC8AF4B72319 CRC64;
 CC
 CC Query Match 76.9%; Score 30; DB 1; Length 462;
 CC Best Local Similarity 57.1%; Pred. No. 58;
 CC Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 YDSNNV 8
 CC Db 87 YDSNNI 93
 CC
 CC RESULT 15
 CC VP87 NPVOP STANDARD; PRT; 624 AA.
 CC ID VP87 NPVOP
 CC AC P17930;
 CC DT 01-NOV-1990 (Rel. 16, Created)
 CC DT 01-NOV-1990 (Rel. 16, Last sequence update)
 CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
 CC DE Capsid protein P87.
 CC GN P87.
 CC OS Oryza pseudoturgata multicausid polynucleosid virus (OPMVV).
 CC OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC OC Nucleopolydiovirus.
 CC OX NCBI_TaxID=164623;
 CC OX [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=90232722; PubMed=2184573;
 CC RA Mueller R., Pearson M.N., Ruseell R.L.Q., Rohmann G.F.;
 CC RT "A capsid-associated protein of the multicausid nuclear polynucleosid
 CC virus of Oryza pseudoturgata: genetic location, sequence,
 CC RT transcriptional mapping, and immunocytochemical characterization."
 CC RL Virology 176:133-144(1990).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=97271300; PubMed=9126251;
 CC RA Ahrens C.H., Ruseell R.R., Funk C.J., Evans J., Harwood S.,
 CC RA Rohmann G.F.;
 CC RT "The sequence of the Oryza pseudoturgata multinuclencapsid nuclear
 CC polydencosid virus genome."
 CC RL Virology 229:381-399(1997).
 CC -1- SUBCELLULAR LOCATION: CAPSID-ASSOCIATED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D13959; BAA03061.1; --
 CC DR EMBL: U75930; AAC59104.1; --
 CC DR PIR: A34602; VCNV87.
 CC KW Coat protein; Repeat.
 CC FT DOMAIN 256 280 2 X 13 AA TANDEM REPEATS, MOTIF A.
 CC FT REPEAT 256 268
 CC FT REPEAT 269 280 MOTIF A.
 CC FT DOMAIN 283 324 3 X 11 AA REPEATS, MOTIF B.
 CC FT REPEAT 283 293 MOTIF B.
 CC FT REPEAT 296 306 MOTIF B.
 CC FT REPEAT 314 324 MOTIF B.
 CC SQ SEQUENCE 624 AA; 70696 MW; 8573D00964C9157B CRC64;
 CC
 CC Query Match 76.9%; Score 30; DB 1; Length 624;
 CC Best Local Similarity 62.5%; Pred. No. 81;

Matches	5;	Conservative	3;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	SYDSSNVV	8						
		: :							
Db	352	SYDTSSVI	359						

Search completed: January 12, 2004, 06:59:48
Job time : 7.375 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:04 ; Search time 16.25 Seconds

(without alignments)
127,041 Million cell updates/sec

Title: US-09-829-495-66

Perfect score: 39
Sequence: 1 SYDSNNV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp Unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archae:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	505	16	Q92F18
2	33	84.6	600	5	Q81TJ3
3	33	84.6	680	10	P931658
4	33	84.6	730	10	Q92U23
5	33	84.6	934	16	Q9PMF2
6	33	84.6	1505	5	Q81M99
7	33	84.6	1945	5	Q96681
8	33	84.6	1945	5	Q9V491
9	33	82.1	238	12	Q93J53
10	32	82.1	507	12	Q8QHV1
11	32	82.1	521	5	Q9VDO3
12	32	82.1	523	5	Q8MT85
13	32	82.1	527	5	P91310
14	32	82.1	542	5	Q9V9K0
15	32	82.1	557	5	Q95XY3
16	32	82.1	571	2	Q54121

17	32	82.1	584	10	Q8H547
18	32	82.1	1938	2	P70983
19	32	82.1	4654	10	Q8D418
20	32	82.1	7149	5	Q81M09
21	31	79.5	158	5	Q19343
22	31	79.5	192	5	Q9V8A2
23	31	79.5	249	10	Q9SUQ4
24	31	79.5	261	5	Q18030
25	31	79.5	274	5	Q17086
26	31	79.5	304	10	Q98RQ1
27	31	79.5	317	5	Q8MR58
28	31	79.5	326	5	Q8WP98
29	31	79.5	333	12	Q8UKJ2
30	31	79.5	358	16	Q97MK4
31	31	79.5	374	5	Q17852
32	31	79.5	393	16	Q8D4N4
33	31	79.5	395	5	Q95WDP
34	31	79.5	398	16	Q9A1Z8
35	31	79.5	398	16	Q8P318
36	31	79.5	400	2	Q08342
37	31	79.5	411	16	Q69522
38	31	79.5	413	16	Q69679
39	31	79.5	414	5	Q963G5
40	31	79.5	447	2	Q9AKA4
41	31	79.5	447	16	Q8E7X9
42	31	79.5	447	16	Q8E2H1
43	31	79.5	450	5	Q17854
44	31	79.5	491	5	Q963G9
45	31	79.5	514	3	Q9Y893

ALIGNMENTS

RESULT 1	Q92F18	PRELIMINARY;	PRT;	505 AA.
AC	Q92F18;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 23, Last sequence update)			
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Internalin like protein (LPGTG motif).			
GN	LIN0290.			
OS	Listeria innocua.			
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.			
OX	NCBI_TaxID=1642;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CLIP 11262 / Serovar 6a;			
RX	PubMed=11679669;			
RA	Glaser P., Frangul L., Buchrieser C., Rusniok C., Amend A.,			
RA	Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,			
RA	Charbit A., Chetoui F., Couve E., de Daruvar A., Deloux P.,			
RA	Dornann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,			
RA	Eutlian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,			
RA	Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,			
RA	Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunat F., Kurpjak G.,			
RA	Madeno E., Maitouram A., Mata Vicente J., Ng E., Nedjati H.,			
RA	Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purrell R.,			
RA	Rommel B., Rose M., Schlueter T., Simoes N., Tietz A.,			
RA	Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,			
RT	"Comparative genomics of Listeria species."			
RL	Science 294:849-852(2001).			
DR	EMBL; AL596164; CAC95523.1; -.			
DR	ListLst; LIN00290; -.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	InterPro; IPR001612; LPGTG.			
DR	InterPro; IPR001611; LRR.			
DR	InterPro; IPR007092; LRR_SDS22.			
DR	Pfam; PF00746; Gram_pos_anchor; 1.			
DR	Pfam; PF00560; LRR; 7.			
DR	TIGRFAMs; TIGR01167; LPGTG_anchor; 1.			
DR	PROSITE; P50847; GRAM_POS_ANCHORING; 1.			

DR PROSITE; PS50504; LRR_SDS22; 3.
 KW Complete proteome.
 SQ SEQUENCE 505 AA; 54909 MW; BCE8B8477E3678CC CRC64;

Query Match 84.6%; Score 33; DB 16; Length 505;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
 DB 319 YDSSNVV 326

RESULT 2

ID Q8ITJ3 PRELIMINARY; PRT; 600 AA.
 AC Q8ITJ3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Glutamylyl transpeptidase.
 OS Heterodera glycines (soybean cyst nematode).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
 NCBI_Taxid=51029;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Gao B., Maier T., Davis E.L., Baum T.J., Husey R.S.;
 RT "Cloning of Heterodera glycines glutamyl transpeptidase gene Hg-
 9C."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF469056; AAN32885.1; -
 SQ SEQUENCE 600 AA; 66232 MW; 67069FEC124453 CRC64;

Query Match 84.6%; Score 33; DB 5; Length 600;
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
 DB 564 SYDSSNVV 571

RESULT 3

ID P93658 PRELIMINARY; PRT; 680 AA.
 AC P93658;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Usamonnate inducible protein.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Brassica.
 NCBI_Taxid=3708;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Global;
 RA Geshi N., Brandt A.;
 RT "Two jasmonate inducible proteins from Brassica napus seedlings
 RT homologous to myrosinase binding proteins and jacalin";
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y11483; CAA72271.1; -
 DR HSSP; P18670; IUAC.
 DR InterPro; IPR001229; Jacalin_lectin.
 DR Pfam; PF01419; Jacalin; 4.
 SQ SEQUENCE 680 AA; 69937 MW; 8CD0CF16C30CA1E2 CRC64;

Query Match 84.6%; Score 33; DB 10; Length 680;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNVV 8
 DB 308 YDSSNVV 314

RESULT 4

ID Q9ZU23 PRELIMINARY; PRT; 730 AA.
 AC Q9ZU23;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE F5F19.6 protein.
 GN F5F19.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsie.
 NCBI_Taxid=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Vysotskaya V.S., Schwartz J.R., Tortum M., Yu G., Lenz C., Liu S.,
 RA Li J., Kremenetskaia I., Luros J., Gonzalez A., Alrafi H., Araujo R.,
 RA Buehler E., Conway A.B., Conn L., Dunn P., Hansen N., Huizar L.,
 RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
 RA Becker J.R., Federapfel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F5F19 sequence."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006216; ABD12691.1; -
 DR HSSP; P18670; IUAC.
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR001229; Jacalin_lectin.
 DR Pfam; PF01419; Jacalin; 3.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
 SQ SEQUENCE 730 AA; 73939 MW; E4333AF93B1D95E CRC64;

Query Match 84.6%; Score 33; DB 10; Length 730;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNVV 8
 DB 511 YDSSNVV 517

RESULT 5

ID Q9PMF2 PRELIMINARY; PRT; 934 AA.
 AC Q9PMF2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative formate dehydrogenase large subunit (Selenocysteine
 containing) (EC 1.2.1.2).
 GN FDHA OR CJ1511C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 NCBI_Taxid=197;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kell J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences";
 RL Nature 403:665-668 (2000).
 DR EMBL; AL139078; CAB73932.1; -
 DR HSSP; P07658; IFDO.

Query Match 84.6%; Score 33; DB 10; Length 934;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DR InterPro: IPR006656; MolYdbOpterin.
 DR InterPro: IPR006963; MolYdbOpt_F44S4.
 DR InterPro: IPR006657; Mol_dinuc_bind.
 DR InterPro: IPR006655; Prok_Moxord.
 DR InterPro: IPR006311; Tat_Moxord.
 DR Pfam: PF00384; molYdbOpterin; 1.
 DR Pfam: PF04879; MolYdbOpt_F44S4; 1.
 DR Pfam: PF01568; MolYdbOpt_binding; 1.
 DR TIGRFAMs: TIGR01409; Tat_signal_seg; 1.
 DR PROSITE: PS00551; MOLYDBOPTERIN_PROK_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 934 AA; 103906 MW; 5613822028D643DA CRC64;
 Query Match 84.6%; Score 33; DB 16; Length 934;
 Best Local Similarity 62.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SYDSNNV 8
 DB 714 SYDSNNI 721
 RESULT 6
 ID Q8IM99 PRELIMINARY; PRT; 1505 AA.
 AC Q8IM99;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG11081-PB.
 GN PLEXA.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
 RA Abril J.F., Agayari A., An H.U., Andrews-Pfankuch C., Baldwin D.,
 RA Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dudson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fioder C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.H., Ibeggam C.,
 RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mikhlov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon S., Nussekm D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banson J., An H., Baldwin D., Banson J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibeggam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mista S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clump M., Drysdale R., Emert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB003846; AA006564.1;
 SQ SEQUENCE 1505 AA; 170234 MW; 0C0D94AD26D9CB5 CRC64;
 Query Match 84.6%; Score 33; DB 5; Length 1505;
 Best Local Similarity 62.5%; Pred. No. 3.5e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYDSNNV 8
 DB 826 SYDSNNI 833
 RESULT 7
 ID Q96681 PRELIMINARY; PRT; 1945 AA.
 AC Q96681;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PLEXA A.
 GN PLEXA OR CG11081.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99091049; PubMed=9875845;
 RA Winberg M.L., Noordermeer J.N., Tamagnone L., Comoglio P.M.,
 RA Spriggs W.K., Tessier-Lavigne M., Goodman C.S.;
 RT "Plexin A is a neuronal semaphorin receptor that controls axon guidance.";

```

RL Cell 95:903-916(1998).
DR EMBL; AF106932; AAD09425.1; -.
DR FlyBase; FBgn0025741; PlexA.
DR InterPro; IPR002909; IPT_TIG.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF01437; PSI; 3.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01833; TIG; 3.
DR SMART; SM00429; IPT; 3.
DR SMART; SM00423; PSI; 3.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00284; SERPIN; 1.
SQ SEQUENCE 1945 AA; 218278 MW; 09D35655980BDEF1 CRC64;

Query Match
Best Local Similarity 84.6%; Score 33; DB 5; Length 1945;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSNNV 8
DB 1266 SYESSNII 1273

RESULT 8
ID Q9V491 PRELIMINARY; PRT; 1945 AA.
AC Q9V491;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PLEXA protein.
GN PLEXA OR CG11081.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
OX NCB1
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bernos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evans G.L., Ferraz C., Ferrer A., Fleischmann W.,
RA Fogle C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodali S.C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

```

RA Spier E., Spradling A.C., Scapleton M., Strong R., Sun E.,
RA Switzek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003846; AAF59394.1; -.
DR FlyBase; FBgn0025741; PlexA.
DR InterPro; IPR002909; IPT_TIG.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF01437; PSI; 3.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01833; TIG; 3.
DR SMART; SM00429; IPT; 3.
DR SMART; SM00423; PSI; 3.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00284; SERPIN; 1.
SQ SEQUENCE 1945 AA; 218356 MW; 4C032CA07C83A745 CRC64;

Query Match
Best Local Similarity 84.6%; Score 33; DB 5; Length 1945;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSNNV 8
DB 1266 SYESSNII 1273

RESULT 9
ID Q99353 PRELIMINARY; PRT; 238 AA.
AC Q99353;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Coat protein (capsid protein).
GN Banana mild mosaic virus.
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Foveavirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Foveavirus.
NCBI_TaxID=148879;
OX NCB1
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21440425; PubMed=11556712;
RA Gambley C.F., Thomas J.E.;
RX "Molecular characterisation of Banana mild mosaic virus, a new
RT filamentous virus in Musa spp.",
RL Arch. Virol. 146:1369-1379(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Gambley C.F., Thomas J.E.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SELF-ASSEMBLES WITH THE RNA TO FORM INFECTIOUS PARTICLES
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE POTEXVIRUSES COAT PROTEIN FAMILY.
DR EMBL; AF314662; AKK28493.1; -.
DR InterPro; IPR000052; P1vix_coat.
DR Pfam; PF00286; virus_P-coat; 1.
DR PRINTS; PR00232; POTXCARLCOAT.
KW Coat protein.
SQ SEQUENCE 238 AA; 26825 MW; 38E7599DB1CDE27A CRC64;

Query Match
Best Local Similarity 82.1%; Score 32; DB 12; Length 238;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNNV 8
DB 31 YDSSNNV 37

```



```

RESULT 10
08OHV1 PRELIMINARY; PRT; 507 AA.
AC 08OHV1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
OC Polypeptide (Fragment).
OS Spartina molle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae.
NCBI_TaxID=190503;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Asasi;
RX MEDLINE=21867386; PubMed=1190529;
RA Goetz R., Hutch W., Lesemann D.E., Mates E.;
RT "Molecular and serological relationships of Spartina molle virus
RT (SPMV) strains from Spartina spec. and from Cynodon dactylon to other
RT members of the Polyviridae."
RT Arch. Virol. 147:379-391(2002).
DR EMBL; AF491352; AAM12550.1; -.
DR InterPro; IPR001592; Poly_coat.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00767; Poly_coat; 1.
DR PROSITE; PS50507; RDRP_POSITIVE; 1.
KM Coat protein.
FT NON_TER
SQ SEQUENCE 507 AA; 57114 MW; DA8184CTDC94CA31 CRC64;

Query Match 82.1%; Score 32; DB 12; Length 507;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
Db 57 SYDSSNV 64

RESULT 11
09VDQ3 PRELIMINARY; PRT; 521 AA.
AC 09VDQ3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG4936 protein (LD08906P).
GN CG4936.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams W.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

```

```

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Moutlow G., Mshelishvili N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weissbrock G.M., Weissbrock J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungali C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouenavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;
RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
RL EMBL; AE003728; AAF5737.1; -.
DR EMBL; AY061100; AAL28648.1; -.
DR Flybase; FBgn038768; CG4936.
DR InterPro; IPR003006; 1g_MHC.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00096; zf-C2H2; 5.
DR SMART; SM00355; ZnF_C2H2; 5.
DR PROSITE; PS00290; 1g_MHC; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 5.
KM Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 521 AA; 58753 MW; 195C4FA4B1830E1F CRC64;

Query Match 82.1%; Score 32; DB 5; Length 521;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
Db 510 SYDSSNV 517

RESULT 12
08MT85 PRELIMINARY; PRT; 523 AA.
AC 08MT85;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GM13256p.
GN CG7861.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnak D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommler B., Li P., Liao G.,

```

RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceinkner S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY118314; AAM48343.1; -.
DR FlyBase; FBgn0033055; CG7861.
DR InterPro; IPR000938; CAP-Gly.
DR Pfam; PF01302; CAP_GLY; 1.
DR PROSITE; PS50245; CAP_GLY_2; 1.
SQ SEQUENCE 523 AA; 55393 MW; 4B75F00038C31F89 CRC64;

Query Match 82.1%; Score 32; DB 5; Length 523;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
Db 95 NYDSSNV 101

RESULT 13

PRELIMINARY; PRT; 527 AA.
ID P91310
AC P91310;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 60.9 kDa protein.
GN F47F6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).

[2]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX "Wilson R., Favello A., Maggi L., Chiapelli B.,
RT "The sequence of C. elegans cosmid F47F6".
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX "Waterston R.;
RT "Direct Submission";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80838; AAC7112.1; -.
DR WormPep; F47F6.3; CE10708.
DR InterPro; IPR002656; Acyl_transf_3.
DR Pfam; PF01757; Acyl_transf_3; 1.
KW Hypothetical protein.
SQ SEQUENCE 527 AA; 60919 MW; 1E459340CE9B6A72 CRC64;

Query Match 82.1%; Score 32; DB 5; Length 527;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
Db 337 SYDSSNV 344

RESULT 14
Q9V9KO PRELIMINARY; PRT; 542 AA.
ID Q9V9KO
AC Q9V9KO;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CG7861 protein.
GN CG7861.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt P.A., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.D., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Misha N.V., Modarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskeir D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Spier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Snie E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Maesman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

DR EMBL; AE003784; AAF57288.2; -.
DR FlyBase; FBgn0033055; CG7861.
DR InterPro; IPR000938; CAP-Gly.
DR Pfam; PF01302; CAP_GLY; 1.
DR Pfam; PF00560; LRR; 1.
DR PROSITE; PS50245; CAP_GLY_2; 1.
SQ SEQUENCE 542 AA; 61580 MW; 6D946885BEA57B5 CRC64;

Query Match 82.1%; Score 32; DB 5; Length 542;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
Db 95 NYDSSNV 101

RESULT 15
Q95XY3 PRELIMINARY; PRT; 557 AA.
ID Q95XY3

AC Q95X13;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 64.1 kDa protein.
GN Y51H7C.12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Bradshaw-Cordum H.; DuBue T.;
RT "The sequence of C. elegans cosmid Y51H7C";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024805; AAK39343.1; -;
DR WormPep; Y51H7C.12; CE26148.
DR InterPro; IPR005069; DUF271.
DR Pfam; PR03407; DUF271.1.
KW Hypothetical protein.
SQ SEQUENCE 557 AA; 64096 MW; A84D0CEAF9D247FF CRC64;
QY 1 SYDSSNV 7
Db 257 SYDSTNV 263

Search completed: January 12, 2004, 07:02:17
Job time : 20.25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:09 ; Search time 19.375 Seconds
(without alignments)
65.539 Million cell updates/sec

Title: US-09-829-495-66

Perfect score: 39

Sequence: 1 SYDSNNV 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	39	100.0	8	22	AA61298
2	39	100.0	8	24	ABU1264
3	39	100.0	161	22	AA61264
4	39	100.0	161	22	AA61264
5	33	84.6	402	15	AA60578
6	33	84.6	730	21	AA629582
7	33	84.6	752	21	AA629581
8	32	82.1	1945	22	AB65449
9	32	82.1	110	23	AA018434
			113	22	AAU02522

10	32	82.1	166	22	AAE09787
11	32	82.1	174	22	AAE09790
12	32	82.1	363	22	AAE09791
13	32	82.1	395	22	AAE09784
14	32	82.1	521	22	AB61042
15	32	82.1	542	22	AB65491
16	32	82.1	798	20	AAW73552
17	32	82.1	893	17	AAW09257
18	32	82.1	1938	17	AAW09255
19	32	82.1	1938	20	AAW73553
20	31	79.5	98	24	ABP5510
21	31	79.5	131	22	AB662470
22	31	79.5	192	22	AB669449
23	31	79.5	220	14	AA42340
24	31	79.5	245	10	AA42340
25	31	79.5	272	22	AB66563
26	31	79.5	325	20	AB602465
27	31	79.5	395	23	ABP25918
28	31	79.5	447	23	ABP25919
29	31	79.5	469	12	AA15510
30	31	79.5	595	23	ABJ05497
31	31	79.5	783	23	AB62054
32	31	79.5	862	23	ABP35634
33	31	79.5	876	23	ABP35634
34	31	79.5	907	22	AB662889
35	31	79.5	964	18	AAW32619
36	31	79.5	1627	23	ABP98127
37	31	79.5	18	24	ABU03437
38	30	76.9	29	13	AA622106
39	30	76.9	126	19	AA645428
40	30	76.9	284	22	ABG16400
41	30	76.9	308	21	AA649677
42	30	76.9	341	22	AAU68573
43	30	76.9	352	22	AB658931
44	30	76.9	354	23	AB692259
45	30	76.9			

ALIGNMENTS

RESULT 1					
ID	AA61298	standard; Peptide; 8 AA.			
XX					
AC	AA61298;				
XX					
DT	04-APR-2001 (first entry)				
XX					
DE	Anti-TANGO 268 scFv CDR, SEQ ID NO: 66.				
XX					
KW	Human; antibody; scFv; CDR; complementarity determining region;				Paenibacillus pabu
KW	TANGO 268; cardiant; cerebrotropic; cytostatic; antitoxigenic;				Paenibacillus pabu
KW	thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI;				Paenibacillus pabu
KW	platelet membrane glycoprotein receptor; bleeding disorder;				Drosophila melanog
KW	blood vessel injury; thrombotic disorder; hemorrhagic disorder;				Drosophila melanog
KW	ischaemia; cardiovascular disease; immunological disease; liver disorder;				Drosophila melanog
KW	cancer.				Drosophila melanog
XX					
OS	Homo sapiens.				
XX					
PN	WO200100810-A1.				
XX					
PD	04-JAN-2001.				
XX					
PF	30-JUN-2000; 2000WO-US18152.				
XX					
PR	30-JUN-1999; 99US-0345468.				
XX					
PR	06-DEC-1999; 99US-0454824.				
XX					
PR	14-FEB-2000; 2000US-0503387.				
XX					
PA	(MILL-) MILLENNIUM PHARM INC.				

PI Busfield SJ, Villedal J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI Qian DM, Kingsbury G;
XX
XX
DR MPI; 2001-080877/09.
XX
XX
PT New genes encoding human platelet-expressed collagen receptor,
PT glycoprotein VI, and its modulators, useful for preventing, treating
PT and diagnosing hemorrhagic disorders, thrombotic diseases and
PT immunological disorders -
XX
XX
PS Claim 32; Page 102; 227pp; English.
XX
XX
CC The present sequence is given in a specification relating to an isolated
CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
CC and polypeptides and their modulators, e.g. antisense nucleic acids,
CC ribozymes and antibodies, are useful for preventing, treating and
CC diagnosing disorders associated with aberrant expression or activity of
CC GPVI. These disorders include bleeding disorders
CC (e.g. thrombocytopenia), blood vessel injury, thrombotic disorders
CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC infarction), immunological diseases (e.g. platelet disorder) and
CC embryonic liver disorders. Preferably they are used to prevent acute
CC cardiac ischaemia following angioplasty and metastatic cancers,
CC especially of the colon and liver.
XX
XX
SQ Sequence 8 AA;
XX
XX
Query Match 100.0%; Score 39; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SYDSNNV 8
Db 1 SYDSNNV 8
XX
XX
RESULT 2
ABU1264
ID ABU1264 standard; Peptide; 8 AA.
XX
XX
AC ABU1264;
XX
XX
DT 06-FEB-2003 (first entry)
XX
XX
DE Human TANGO 268 VLCDRI Peptide #3.
XX
XX
KM Human; mouse; variable heavy; VH; antigen; cancer;
KM complementarity determining region; TANGO 268; glycoprotein VI; GPVI;
KM TANGO 268; extracellular matrix; collagen; platelet release;
KM proliferation; migration; embryogenesis; inflammation; thrombosis;
KM degradation; thrombocytopenia; antibody; thrombotic disorder;
KM cerebral vascular disease; stroke; ischaemia; venous thromboembolism;
KM leg swelling; pain; ulceration; pulmonary embolism; coronary disease;
KM cardiovascular disease; angina pectoris; myocardial infarction;
KM coronary restenosis; atherosclerosis; immunological disorder;
KM developmental disorder; embryonic disorder; liver disorder;
KM cerebral vascular disease; venous thromboembolism disease.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200280968-A1.
XX
XX
PD 17-OCT-2002.
XX
XX
PF 09-APR-2002; 2002WO-US11122.
XX
XX
PR 09-APR-2001; 2001US-0829495.
XX
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX

PI Busfield SJ, Villedal J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI Qian DM, Kingsbury G;
XX
XX
DR MPI; 2003-058477/05.
XX
XX
PT Novel substantially purified antibody immunospecifically binding to
PT TANGO 268 antigen, useful for treating bleeding disorders such as
PT thrombocytopenia, stroke, ischaemia, pulmonary embolism, atherosclerosis
PT -
XX
XX
PS Claim 12; Page 111; 236pp; English.
XX
XX
CC This invention relates to a novel purified antibody comprising a
CC variable heavy (VH) complementarity determining region (CDR)1, VH CDR2
CC or VH CDR3; or variable light (VL) CDR1, VL CDR2 or VL CDR3, and
CC immunospecifically binding to a TANGO 268 (also referred as glycoprotein
CC VI (GPVI)) antigen. The antibodies of the invention act to decrease or
CC block TANGO 268 binding to extracellular matrix components, or as a
CC collagen or platelet release and aggregation blocker. The antibodies of
CC the invention are useful for modulating proliferation, migration,
CC morphology, differentiation and/or function of megakaryocytes and
CC platelets, including during development e.g. embryogenesis, modulating
CC leukocyte-platelet and platelet-endothelium interactions in
CC inflammation and/or thrombosis, and modulating platelet aggregation and
CC degradation. They are also useful for modulating disorders associated
CC with abnormal or aberrant megakaryocyte and/or platelet proliferation,
CC migration, morphology, differentiation and/or function, e.g. bleeding
CC disorders such as thrombocytopenia. Other diseases which may be
CC modulated by these antibodies are thrombotic disorders, cerebral
CC vascular diseases (e.g. stroke and ischaemia) venous thromboembolism
CC diseases (e.g. diseases involving leg swelling, pain and ulceration,
CC pulmonary embolism, etc); coronary diseases (e.g. cardiovascular
CC diseases including angina pectoris, myocardial infarction, coronary
CC restenosis, atherosclerosis, etc); immunological disorders,
CC developmental disorders, embryonic disorders, liver disorders, cerebral
CC vascular diseases, venous thromboembolism disease, coronary diseases,
CC and metastatic cancers. The antibodies of the invention only causes a
CC transient decrease in platelet counts, platelet aggregation, and/or
CC platelet activation and so have some advantages over prior art
CC methods. The present sequence represents a peptide sequence used to
CC generate the antibodies of the invention.
XX
XX
SQ Sequence 8 AA;
XX
XX
Query Match 100.0%; Score 39; DB 24; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SYDSNNV 8
Db 1 SYDSNNV 8
XX
XX
RESULT 3
AAG76011
ID AAG76011 standard; Protein; 161 AA.
XX
XX
AC AAG76011;
XX
XX
DT 03-SEP-2001 (first entry)
XX
XX
DE Human colon cancer antigen protein SEQ ID NO:6775.
XX
XX
KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; chromosome 22.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200122920-A2.
XX
XX
PD 05-APR-2001.
XX
XX
PF 28-SEP-2000; 2000WO-US26524.
XX

XX	29-SEP-1999;	99US-0157137.	
PR	03-NOV-1999;	99US-0163280.	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Ruben SM, Barash SC, Birse CE, Rosen CA;		
XX			
DR	WPI, 2001-235357/24.		
XX	N-PSDB; AAH35416.		
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,		
PS	useful for preventing, diagnosing and/or treating colorectal cancers -		
XX	Claim 11, Page 8234-8235; 9803JP, English.		
XX			
CC	AAH32943 to AAH31195 and AAG73514 to AAG77788 represent human colon		
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where		
CC	the proteins are collectively known as colon cancer antigens. The colon		
CC	cancer antigens have cytosolic activity and can be used in gene		
CC	therapy and vaccine production. N and P may be used in the prevention,		
CC	diagnosis and treatment of diseases associated with inappropriate P		
CC	expression. For example, N and P may be used to treat disorders		
CC	associated with decreased expression by rectifying mutations or deletions		
CC	in a patient's genome that affect the activity of P by expressing		
CC	inactive proteins or to supplement the patient's own production of P.		
CC	Additionally, N may be used to produce the colon cancer-associated Ps,		
CC	by inserting the nucleic acids into a host cell and culturing the cell		
CC	to express the proteins. N and P can be used in the prevention, diagnosis		
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204		
CC	and AAB77789 represent sequences used in the exemplification of the		
CC	present invention.		
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were		
CC	missing at time of publication, meaning no sequences are present for		
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.		
XX			
SQ	Sequence	161 AA;	
	Query Match.	100.0%; Score 39; DB 22; Length 161;	
	Best Local Similarity	100.0%; Pred. No. 1.9;	
	Matches	8; Conservative	0; Mismatches
		0; Indels	0; Gaps
OY	1 SYDSSNV 8		
	114 SYDSSNV 121		
DB			
	RESULT 4		
ID	AAH60578	standard; Protein; 402 AA.	
XX	AAH60578;		
AC			
XX			
DT	25-MAR-2003	(updated)	
DT	30-MAR-1995	(first entry)	
XX			
DE	Osteogenic protein OP2.		
XX			
KM	Osteogenic protein.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	Cleavage-site	239..242	
FT	Cleavage-site	/note="proteolytic cleavage site"	
FT		260..263	
FT		/note="proteolytic cleavage site"	
XX			
PN	WO9420539-A1.		
XX			
PD	15-SEP-1994.		
XX			
PF	04-MAR-1994;	94WO-US02335.	

PR	04-MAR-1993;	93US-0027070.	
XX			
PA	(CREA-) CREATIVE BIOMOLECULES INC.		
XX			
PI	Jones WK, Oppermann H, Ozkaynak E, Rueger DC, Sampath KT,		
PI	Tucker RF;		
XX			
DR	WPI, 1994-302971/37.		
DR	N-PSDB; AAQ71426.		
XX			
PT	Binding partners, esp. antibodies, specific for different forms		
PT	of osteogenic protein - for differentiating between mature and		
PT	soluble complexed forms of the protein in culture media or serum.		
XX			
PS	Disclosure; Page 50-52; 70pp; English.		
XX			
CC	The osteogenic protein is produced recombinantly in mammalian cell		
CC	cultures, and may be provided to a site for bone induction in a		
CC	mammal with a suitable matrix to allow infiltration, proliferation		
CC	and differentiation of migrating progenitor cells.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
SO	Sequence	402 AA;	
	Query Match	89.7%; Score 35; DB 15; Length 402;	
	Best Local Similarity	100.0%; Pred. No. 38;	
	Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	2 YDSSNVV 8		
DB	380 YDSSNV 386		
RESULT 5			
AAAG29582			
ID	AAAG29582 standard; Protein: 730 AA.		
XX			
AC	AAAG29582;		
XX			
DT	17-OCT-2000 (first entry)		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 35222.		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	termination sequence.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
XX			
PR	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126284.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	23-APR-1999; 99US-0130891.		
PR	28-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	30-APR-1999; 99US-0132407.		

PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132488.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140635.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141882.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142330.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143562.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 19-JUL-1999; 99US-0144336.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 21-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 84.6%; Score 33; DB 21; Length 730;
 Best Local Similarity 71.4%; Pred. No. 2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSNNV 8
 DB 511 YDSNNI 517

RESULT 6
 AAG29581
 ID AAG29581 standard; Protein; 752 AA.
 AC AAG29581;
 XX
 XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 35221.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.

```

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147203.
PR 04-AUG-1999; 99US-0147208.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147483.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160776.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

```

```

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match          84.6%; Score 33; DB 21; Length 752;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 YDSNVV 8
DB 533 YDSSNII 539

```

RESULT 7

ABB65449 ID ABB65449 standard; Protein; 1945 AA.

XX ABB65449;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 23139.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EM;

DR WPI; 2001-656860/75.

XX N-PSDB; ABL09552.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Disclosure; SEQ ID NO 23139; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (AB16176-AB130511).

CC (AB16176-AB130511) and the encoded proteins

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ffp.wipo.int/pub/published_pct_sequences.

XX Sequence 1945 AA;

```

Query Match          84.6%; Score 33; DB 22; Length 1945;
Best Local Similarity 62.5%; Pred. No. 6.2e+02;

```


XX	28-FEB-2001; 2001WO-DK00132.
XX	01-MAR-2000; 2000DK-0000326.
XX	(NOVO) NOVOZYMES AS.
XX	Wiltling R, Bjornvad ME, Kauppinen MS, Schuelein M;
XX	WPI; 2001-565502/63.
XX	
DR	
XX	
PT	Novel xyloglucanase enzyme for use in textile, detergent and cellulose
PR	fiber processing industries comprises family 5 of glycosyl hydrolases
PT	and is derived from strains of Paenibacillus -
XX	
PS	Example 2; Page 91-92; 99pp; English.
XX	
CC	The present sequence is Paenibacillus pabuli xyloglucanase XYG1035 enzyme
CC	C-terminal sequence which belongs to the family 5 of glycosyl hydrolase.
CC	Xyloglucanase is useful for machine treatment of fabrics during a washing
CC	cycle of a machine washing process with a washing solution. Xyloglucanase
CC	is useful in textile industry for improving the properties of cellulosic
CC	fibres, yarn, woven or non-woven fabric and in scouring process step, in
CC	cellulose fibre processing industry for ratfing of fibres such as hemp,
CC	jute, flax and linen. Xyloglucanase is useful for treatment of cellulosic
CC	material during the processing of cellulosic material into a material
CC	ready for garment manufacture or fabric manufacture, e.g. in the desizing
CC	or scouring or during industrial or household laundering of such fabric
CC	or garment. Xyloglucanase is also useful in detergent compositions for
CC	removing or bleaching certain soils or stains present on laundry,
CC	especially soils and spots resulting from xyloglucan-containing food and
CC	plants. Xyloglucanase prevents binding of certain soils to the xyloglucan
CC	left on the cellulosic material and shows high performance in
CC	conventional liquid detergent compositions.
SQ	Sequence 166 AA;
OY	Query Match 82.1%; Score 32; DB 22; Length 166;
	Best Local Similarity 87.5%; Pred.No. 60;
	Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
DB	1 SYDSSNVV 8 108 SYDSSNNV 115
RESULT 11	
ID	AAB09790 AAE09790 standard; Protein; 174 AA.
AC	AAB09790;
DT	29-NOV-2001 (first entry)
DE	Paenibacillus pabuli xyloglucanase XYG1037 enzyme C-terminal sequence.
KM	Xyloglucanase; XYG1037; glycosyl hydrolase; textile industry; detergent;
KW	cellulose fibre; hemp; jute; flax; linen; garment manufacture.
OS	Paenibacillus pabuli.
PN	WO200164853-A1.
PD	07-SEP-2001.
PF	28-FEB-2001; 2001WO-DK00132.
PR	01-MAR-2000; 2000DK-0000326.
PA	(NOVO) NOVOZYMES AS.
PI	Wiltling R, Bjornvad ME, Kauppinen MS, Schuelein M;
WI	WPI; 2001-565502/63.

```

PT  Novel xyloglucanase enzyme for use in textile, detergent and cellulose
PT  fiber processing industries comprises family 5 of glycosyl hydrolases
PT  and is derived from strains of Paenibacillus
PT
XX
PS  Example 2; Page 93; 99pp; English.
XX
CC  The present sequence is Paenibacillus pabuli xyloglucanase XYG1037 enzyme
CC  C-terminal sequence which belongs to the family 5 of glycosyl hydrolase.
CC  Xyloglucanase is useful for machine treatment of fabrics during a washing
CC  cycle of a machine washing process with a washing solution. Xyloglucanase
CC  is useful in textile industry for improving the properties of cellulosic
CC  fibres, yarn, woven or non-woven fabric and in scouring process step, in
CC  cellulose fibre processing industry for retting of fibres such as hemp,
CC  jute, flax and linen. Xyloglucanase is useful for treatment of cellulosic
CC  material during the processing of cellulosic material into a material
CC  ready for garment manufacture or fabric manufacture, e.g. in the desizing
CC  or scouring or during industrial or household laundering of such fabric
CC  or garment. Xyloglucanase is also useful in detergent compositions for
CC  removing or bleaching certain soils or stains present on laundry,
CC  especially soils and spots resulting from xyloglucan-containing food and
CC  plants. Xyloglucanase prevents binding of certain soils to the xyloglucan
CC  left on the cellulosic material and shows high performance in
CC  conventional liquid detergent compositions.
CC
XX  Sequence 174 AA;
XX
SQ
XX
CY  Query Match 82.1%; Score 32; DB 22; Length 174;
XX  Best Local Similarity 87.5%; Pred. No. 64;
XX  Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 SYDSSNNV 8
XX 117 SYDSSNNV 124
XX
DB
XX
RESULT 12
XX  AAE09791
XX  AAE09791 standard; Protein; 363 AA.
XX
XX  AAE09791;
XX
XX 29-NOV-2001 (first entry)
XX
XX Paenibacillus pabuli mature xyloglucanase XYG1034 enzyme.
XX
XX DE Paenibacillus pabuli mature xyloglucanase XYG1034 enzyme.
XX
XX KM Xyloglucanase; XYG1034; glycosyl hydrolase; textile industry; detergent;
XX cellulose fibre; hemp; jute; flax; linen; garment manufacture.
XX
XX OS Paenibacillus pabuli.
XX
XX PN WO200164853-A1.
XX
XX PD 07-SEP-2001.
XX
XX PF 28-FEB-2001; 2001WO-DK00132.
XX
XX PR 01-MAR-2000; 2000DK-0000326.
XX
XX PA (NOVO ) NOVOZYMES AS.
XX
XX PI Wiltling R, Bjornvad ME, Kauppinen MS, Schuelein M;
XX
XX WPI; 2001-565502/63.
XX
XX DR N-PSDB; AAD16819.
XX
XX Novel xyloglucanase enzyme for use in textile, detergent and cellulose
XX fiber processing industries comprises family 5 of glycosyl hydrolases
XX and is derived from strains of Paenibacillus
XX
XX Example 7; Page 95; 99pp; English.
XX
XX The present sequence is Paenibacillus pabuli xyloglucanase XYG1034
XX

```


CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 521 AA;

Query Match 82.1%; Score 32; DB 22; Length 521;

Best Local Similarity 62.5%; Pred. No. 2.2e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 8

DB 510 SYDTANIV 517

RESULT 15

ABB58491 ID ABB58491 standard; Protein; 542 AA.

XX ABB58491;

AC ABB58491;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 2265.

XX Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li FWD, Myers EW,

PI WPI; 2001-656860/75.

XX N-PSDB; ABL02594.

DR New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Disclosure; SEQ ID NO 2265; 21pp + Sequence Listing; English.

PS The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

Search completed: January 12, 2004, 06:58:59
 Job time : 21.375 secs

XX Sequence 542 AA;

Query Match 82.1%; Score 32; DB 22; Length 542;

Best Local Similarity 85.7%; Pred. No. 2.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7

DB 95 NYDSSNV 101

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:49 ; Search time 6.34375 Seconds
(without alignments)
106.117 Million cell updates/sec

Title: US-09-829-495-65

Perfect score: 38
Sequence: 1 EDNORPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	92.1	112	1 L6HUR	Ig lambda chain V-
2	34	89.5	109	1 L1HWA	Ig lambda chain V-
3	34	89.5	111	1 L1HWA	Ig lambda chain V-
4	34	89.5	131	1 L6HUR	Ig lambda chain pr
5	34	89.5	266	2 S72947	hypothetical prote
6	34	89.5	704	2 T02558	hypothetical prote
7	33	86.8	111	1 L6HUR	Ig lambda chain V-
8	32	84.2	111	1 L6HUR	Ig lambda chain V-
9	32	84.2	460	2 T27770	hypothetical prote
10	32	84.2	471	2 T20690	hypothetical prote
11	32	84.2	603	2 AC0674	probable glycoen
12	31	81.6	107	2 PC4283	anti-SS-A/Ro 60K p
13	31	81.6	161	2 E81447	transcription elon
14	30	78.9	108	1 LSHDL	Ig lambda chain V-
15	30	78.9	158	2 C95159	shikimate kinase
16	30	78.9	158	2 C95159	shikimate kinase
17	30	78.9	232	2 C84706	probable small nuc
18	30	78.9	332	2 A45710	BRF2 gene homolog
19	30	78.9	361	2 T30402	hypothetical prote
20	30	78.9	371	1 E64821	Y111 protein precu
21	30	78.9	371	2 E90743	probable dehydroge
22	30	78.9	420	2 B86217	protein T27G7.4 [i
23	30	78.9	462	1 S35534	adenovirus E1A enh
24	30	78.9	522	2 D82195	phase replication
25	30	78.9	555	1 S24061	transcription fact
26	30	78.9	620	1 S33253	protein-tyrosine k
27	30	78.9	664	2 T01368	hypothetical prote
28	30	78.9	885	1 A55453	nucleotide diphosp
29	30	78.9	885	1 A55453	nucleotide diphosp

30	29	76.3	98	2 S36068	Ig lambda chain -
31	29	76.3	106	2 S38495	Ig lambda chain -
32	29	76.3	111	2 S47009	Ig lambda chain Vi
33	29	76.3	112	2 S51125	genome polypeptide
34	29	76.3	117	2 S04525	Ig lambda chain pr
35	29	76.3	132	2 I39004	cyclin-dependent k
36	29	76.3	195	2 A27241	helix-destabilizin
37	29	76.3	202	1 MFVNRV	matrix protein - r
38	29	76.3	202	1 MFVNAV	matrix protein - r
39	29	76.3	202	1 MFVNAV	matrix protein - r
40	29	76.3	202	1 MFVNAV	matrix protein - r
41	29	76.3	232	2 S25742	Ig lambda chain -
42	29	76.3	243	2 A96744	hypothetical prote
43	29	76.3	312	1 S22397	pyruvate synthase
44	29	76.3	312	2 E84268	hypothetical prote
45	29	76.3	318	2 H69959	hypothetical prote

ALIGNMENTS

RESULT 1
L6HUR
Ig lambda chain V-VI region (AR) - human (tentative sequence)
C/Species: Homo sapiens (man)
C/Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 31-Mar-2000
C/Accession: A01987
R/Sletten, K.; Navtig, J.B.; Hueby, G.; Junl, J.
Biochem. J. 195, 561-572, 1981
A/Title: The complete amino acid sequence of a prototype immunoglobulin-lambda light-chain
A/Reference number: A01987; PMID:82091000; PMID:6797401
A/Contents: amyloid protein AR
A/Accession: A01987
A/Molecule type: protein
A/Residues: 1-112 <SLR>
A/Note: about half of the lambda chain C region is missing from this protein
C/Comment: This protein was isolated from the spleen of a patient with amyloidosis.
C/Genetics:
A/Gene: GDB:IGLV@
A/Cross-references: GDB:119342; OMIM:147240
A/Map position: 22q11.2-22q11.2
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: amyloid; heterotetramer; immunoglobulin
F:15-93/Domain: immunoglobulin homology <IMM>
F:22-91/Disulfide bonds: #status predicted

Query Match 92.1% Score 35; DB 1; Length 112;
Best Local Similarity 85.7% Pred. No. 2.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 DDNORPS 57

RESULT 2
L1HWA
Ig lambda chain V-I region (Wah) - human
C/Species: Homo sapiens (man)
C/Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 02-Sep-1997
C/Accession: A01967
R/Takahashi, Y.; Takahashi, N.; Tetaert, D.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 80, 3686-3690, 1983
A/Title: Complete covalent structure of a human immunoglobulin D: sequence of the lambda A/Reference number: A01967; PMID:8322161; PMID:6407018
A/Accession: A01967
A/Molecule type: protein
A/Residues: 1-109 <TK>
C/Genetics:
A/Gene: GDB:IGLV@
A/Cross-references: GDB:119342; OMIM:147240

A/Map position: 22q11.2-22q11.2
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) heavy disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin
F/1-97/Region: V segment
F/15-91/Domain: immunoglobulin homology <IMM>
F/98-109/Region: J segment
F/22-89/Disulfide bonds: #status predicted

Query Match 89.5%; Score 34; DB 1; Length 109;
Best Local Similarity 85.7%; Pred. No. 3.5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 EDNORPS 57

RESULT 3
L1HUNW
Ig lambda chain V-I region (New) - human
C/Species: Homo sapiens (man)
C/Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 24-Nov-1999
C/Accession: A01964
R/Langer, B.; Steinmetz-Kayne, M.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 349, 945-951, 1968
A/Title: The complete amino acid sequence of Bence-Jones protein New (lambda type). Subg
A/Reference number: A01964; MUID:69050892; PMID:4177823
A/Accession: A01964
A/Molecule type: protein
A/Residues: 1-111 <LAN>
C/Comment: This is a Bence Jones protein.
C/Genetics: GDB:IGLV@
A/Gene: GDB:IGLV@
A/Cross-references: GDB:119342; OMIM:147240
A/Map position: 22q11.2-22q11.2
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) heavy disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: blocked amino end; heterotetramer
F/15-91/Domain: immunoglobulin homology <IMM>
F/1/Modified site: blocked amino end (Gln) (probably pyrrolidone carboxylic acid) #statu
F/22-89/Disulfide bonds: #status predicted

Query Match 89.5%; Score 34; DB 1; Length 111;
Best Local Similarity 85.7%; Pred. No. 3.5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 EDNORPS 57

RESULT 4
L6HUBH
Ig lambda chain precursor V-VI region (EB4) - human
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
C/Accession: A01990
R/Anderson, M.L.M.; Brown, L.; McKenzie, E.; Kellow, J.E.; Young, B.D.
Nucleic Acids Res. 13, 2931-2941, 1985
A/Title: Cloning and sequence analysis of an Ig lambda light chain mRNA expressed in the
A/Reference number: A01990; MUID:95215660; PMID:3923440
A/Accession: A01990
A/Molecule type: mRNA
A/Residues: 1-131 <AND>
C/Genetics: GDB:IGLV@
A/Gene: GDB:IGLV@
A/Cross-references: GDB:119342; OMIM:147240
A/Map position: 22q11.2-22q11.2
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) heavy disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a

C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-131/Product: Ig lambda chain V-VI region (EB4) #status predicted <MAT>
F/20-41/Region: framework 1
F/34-112/Domain: immunoglobulin homology <IMM>
F/42-54/Region: complementarity-determining 1
F/55-69/Region: framework 2
F/70-76/Region: complementarity-determining 2
F/77-110/Region: framework 3
F/111-118/Region: complementarity-determining 3
F/119-131/Region: framework 4
F/41-110/Disulfide bonds: #status predicted

Query Match 89.5%; Score 34; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORP 6
DB 70 EDNORP 75

RESULT 5
S75947
hypothetical protein B2235 F2.77 - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
C/Accession: S75947
R/Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A/Description: Mycobacterium leprae cosmid B2235.
A/Reference number: S72587
A/Accession: S72947
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-266 <SMI>
A/Cross-references: EMBL:U00019; NID:9467079; PIDN:AAA17283.1; PID:9467100

Query Match 89.5%; Score 34; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORP 6
DB 103 EDNORP 108

RESULT 6
T02558
hypothetical protein AC2932590 [imported] - Arabidopsis thaliana
N/Alternate names: hypothetical protein T26B15.15
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 02-Feb-2001
C/Accession: T02558; AB4735
R/Rounseley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, July 1998
A/Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.
A/Reference number: Z14678
A/Accession: T02558
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-704 <ROU>
A/Cross-references: EMBL:AC004681; NID:93298532; PID:93298547
A/Experimental source: cultivar Columbia
R/Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Talion, L. eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; MUID:20083487; PMID:10617197
A/Accession: AB4735
A/Status: preliminary

A:Molecule type: DNA
A:Residues: 1-704 <STO>
A:Cross-References: GB:AE002093; NID:g3298547; PIDN:AAC25941.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32590; T26815.15
A:Map position: 2
A:introns: 89/3; 123/3; 163/1; 184/3; 205/1; 400/3; 449/1; 543/1; 594/3; 631/3

Query Match
Best Local Similarity 89.5%; Score 34; DB 2; Length 704;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 330 ENNORPS 336

RESULT 7
L6HULT
Ig lambda chain V-VI region (WLT) - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
C:Accession: A01989
R:Duillet, F.E.; Strako, K.; Benson, M.D.
Scand, J. Immunol. 22, 653-660, 1985
A:Title: Amino acid sequence of a lambda VI primary (AL) amyloid protein (WLT).
A:Reference number: A01989; MUID:86122667; PMID:4089539
A:Accession: A01989
A:Molecule type: protein
A:Residues: 1-111 <DMU>
C:Genetics:
A:Gene: GDB:IGLV@
A:Cross-References: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as Iga and IgM, the subunits associate into 16 C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-22/Region: framework 1
F:15-93/Domain: immunoglobulin homology <IMM>
F:23-35/Region: complementarity-determining 1
F:36-50/Region: framework 2
F:51-57/Region: complementarity-determining 2
F:58-91/Region: framework 3
F:92-101/Region: complementarity-determining 3
F:102-111/Region: framework 4
F:22-91/Disulfide bonds: #status predicted

Query Match
Best Local Similarity 86.8%; Score 33; DB 1; Length 111;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 ENNORPS 57

RESULT 8
L6HUST
Ig lambda chain V-VI region (SUT) - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
C:Accession: A01988
R:Solomon, A.; Kyle, R.A.; Frangione, B.
In Amyloidosis, Glenner, G.G., Osseman, E.F., Benditt, E.P., Calkins, E., Cohn, A.S.,
A:Title: Light chain variable region subgroups of monoclonal immunoglobulins in amyloid
A:Reference number: A01988
A:Accession: A01988
A:Molecule type: protein
A:Residues: 1-111 <SOL>
C:Genetics:
A:Gene: GDB:IGLV@
A:Cross-References: GDB:119342; OMIM:147240

A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as Iga and IgM, the subunits associate into 16 C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-22/Region: framework 1
F:15-93/Domain: immunoglobulin homology <IMM>
F:23-35/Region: complementarity-determining 1
F:36-50/Region: framework 2
F:51-57/Region: complementarity-determining 2
F:58-91/Region: framework 3
F:92-100/Region: complementarity-determining 3
F:101-111/Region: framework 4
F:22-91/Disulfide bonds: #status predicted

Query Match
Best Local Similarity 84.2%; Score 32; DB 1; Length 111;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 EDORPS 57

RESULT 9
T27770
hypothetical protein ZK177.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T27770
R:Anderson, K.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid ZK177.
A:Reference number: Z20416
A:Accession: T27770
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-460 <AND>
A:Cross-References: EMBL:U21321; PIDN:AAB36965.1; GSPDB:GN00020; CESP:ZK177.1
A:Experimental source: strain Bristol N2; clone ZK177
C:Genetics:
A:Gene: CESP:ZK177.1
A:Map position: 2
A:introns: 20/1; 31/3; 71/3; 111/1; 127/1; 294/1; 323/1
C:Superfamily: Caenorhabditis elegans hypothetical protein ZK177.1

Query Match
Best Local Similarity 84.2%; Score 32; DB 2; Length 460;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 153 QDNORPA 159

RESULT 10
T20690
hypothetical protein F10B5.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T20690
R:Sim, M.
submitted to the EMBL Data Library, February 1995
A:Reference number: Z19310
A:Accession: T20690
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <MIL>
A:Cross-References: EMBL:Z48334; PIDN:CA88310.1; GSPDB:GN00020; CESP:F10B5.3
A:Experimental source: clone F10B5
C:Genetics:
A:Gene: CESP:F10B5.3
A:Map position: 2

A:introns: 36/1; 114/3; 319/1; 362/2

Query Match 84.2%; Score 32; DB 2; Length 471;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 3 EDNORPS 9

RESULT 11

AC0674

probable glycogen debranching protein (EC 3.2.1.-) [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhimurium

C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhimurium

A:Note: this species has also been called *Salmonella typhimurium*

C:Date: 03-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AC0674

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.

Nature 413, 848-852, 2001

A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AC0674

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-603 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD01764.1; PID:G16502612; GSPDB:GN00176

C:Genetics:

A:Gene: SRY1505

C:Superfamily: glyx protein

C:Keywords: glycosidase; hydrolase

Query Match 84.2%; Score 32; DB 2; Length 603;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 154 EDNORPS 160

RESULT 12

PC4283

anti-S8-A/Ro 60K peptide light chain E-60 - human (fragment)

C:Species: *Homo sapiens* (man)

C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 16-Jul-1999

C:Accession: PC4283

R:Suzuki, H.; Takemura, H.; Suzuki, M.; Sekine, Y.; Kashiwagi, H.

Biochem. Biophys. Res. Commun. 237, 101-106, 1997

A:Title: Molecular cloning of anti-S8-A/Ro 60-KDa peptide fab fragments from infiltrating T cells

A:Reference number: PC4279; MUID:97236289; PMID:9125110

A:Accession: PC4283

A:Molecule type: protein

A:Residues: 1-107 <SUZ>

C:Comment: This antibody is commonly found in systemic autoimmune diseases such as Sjogren's syndrome

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:13-87/Domain: immunoglobulin homology <IMM>

Query Match 81.6%; Score 31; DB 2; Length 107;
Best Local Similarity 71.4%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 47 QDNKRS 53

RESULT 13

E81447

transcription elongation factor Cj0287c [imported] - *Campylobacter jejuni* (strain NCTC 11472)

C:Species: *Campylobacter jejuni*

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: E81447

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Farrar, S.; Moule, S.; O'Garra, P.

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariability in surface-associated genes

A:Reference number: AB1250; MUID:20150912; PMID:10688204

A:Accession: E81447

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <PAR>

A:Cross-references: GB:AL139074; GB:AL111168; NID:G6367505; PIDN:CAJ27254.1; PID:G636776

A:Experimental source: serotype O2, strain NCTC 11472

C:Genetics:

A:Gene: greA; Cj0287c

C:Superfamily: transcription elongation factor greB

C:Keywords: transcription factor

Query Match 81.6%; Score 31; DB 2; Length 161;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 22 KDNORPA 28

RESULT 14

LSH0DL

Ig lambda chain V-V region (Del) - human (tentative sequence)

C:Species: *Homo sapiens* (man)

C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000

C:Accession: A01985

R:Butler, M.

Eur. J. Biochem. 50, 49-63, 1974

A:Title: A new subgroup of human L-chains of the lambda-type. Primary structure of Bence-Jones protein

A:Reference number: A01985; MUID:75112179; PMID:4452363

A:Accession: A01985

A:Molecule type: protein

A:Residues: 1-108 <EUJ>

A:Note: This is the first sequenced V region of lambda chain subgroup V

C:Comment: This is a Bence Jones protein.

C:Genetics:

A:Gene: GDB:IGLV@

A:Cross-references: GDB:119342; OMIM:147240

A:Map position: 22q11.2-22q11.2

C:Complex: An immunoglobulin heterodimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterodimer

F:14-68/Domain: immunoglobulin homology <IMM>

F:21-86/Disulfide bonds: #status predicted

Query Match 78.9%; Score 30; DB 1; Length 108;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 48 EDNDRA 54

RESULT 15

C95159

Shikimate kinase [imported] - *Streptococcus pneumoniae* (strain TIGR4)

C:Species: *Streptococcus pneumoniae*

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C:Accession: C95159

R:Petelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Helton, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, M.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

transcription elongation factor Cj0287c [imported] - *Campylobacter jejuni* (strain NCTC 11472)

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
 A:Reference number: A95000; WUID:21357209; PMID:11463916
 A:Accession: C95159
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-158 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75468.1; PID:g14972855; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SPI370
 C:Superfamily: shikimate kinase; shikimate kinase homology

 Query Match 78.9%; Score 30; DB 2; Length 158;
 Best local Similarity 83.3%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 EDNORP 6
 :|||||
 Db 109 KDNORP 114

Search completed: January 12, 2004, 07:03:07
 Job time : 9.34375 secs

THIS PAGE RI ANY (15870)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:49 ; Search time 3.82812 Seconds
(without alignments)

85.992 Million cell updates/sec

Title: US-09-829-495-65

Perfect score: 38

Sequence: 1 EDNORPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	92.1	112	1	LV6A_HUMAN
2	34	89.5	109	1	LV1F_HUMAN
3	34	89.5	111	1	LV1C_HUMAN
4	34	89.5	131	1	LV6E_HUMAN
5	33	86.8	111	1	LV6D_HUMAN
6	33	86.8	494	1	PEA3_BRARE
7	32	84.2	111	1	LV6C_HUMAN
8	32	84.2	460	1	YS41_CAEEL
9	31	81.6	161	1	GREB_CAMJE
10	30	78.9	108	1	LV5A_HUMAN
11	30	78.9	332	1	YIE2_HSVB4
12	30	78.9	371	1	YLTI_ECOLI
13	30	78.9	551	1	ETV4_HUMAN
14	30	78.9	555	1	ETV4_MOUSE
15	30	78.9	620	1	ITX_HUMAN
16	30	78.9	862	1	NPP2_MOUSE
17	30	78.9	885	1	NPP2_RAT
18	29	76.3	195	1	ROA1_BOVIN
19	29	76.3	202	1	VMAT_RABVA
20	29	76.3	202	1	VMAT_RABVC
21	29	76.3	202	1	VMAT_RABVE
22	29	76.3	202	1	VMAT_RABVN
23	29	76.3	202	1	VMAT_RABVP
24	29	76.3	318	1	YCHO_BACST
25	29	76.3	319	1	ROA1_WACMU
26	29	76.3	319	1	ROA1_MOUSE
27	29	76.3	319	1	ROA1_RAT
28	29	76.3	371	1	ROA1_HUMAN
29	29	76.3	473	1	DNPE_MOUSE
30	29	76.3	475	1	PRTG_ERWCH
31	29	76.3	477	1	ETV1_HUMAN
32	29	76.3	477	1	ETV1_MOUSE
33	29	76.3	510	1	ERM_HUMAN

34	29	76.3	519	1	KCG4_HUMAN	Q8tdn1 homo sapien
35	29	76.3	578	1	YV08_SCHPO	O14226 schizosacch
36	29	76.3	871	1	SYA_AQUPY	O9x4dm3 aquilex pyr
37	29	76.3	924	1	ORCI_DROME	O16810 drosophila
38	29	76.3	1131	1	PHYA_SORBI	P93526 sorghum bic
39	29	76.3	1140	1	RA18_SCHRO	P53692 schizosacch
40	29	76.3	1193	1	DPOL_ADEB4	P87503 human adeno
41	29	76.3	1239	1	VI20_EBV	P03189 Epstein-Bar
42	29	76.3	1682	1	CIN7_HUMAN	O01118 homo sapien
43	29	76.3	2164	1	POLG_HRV89	P07210 human rhino
44	29	76.3	5376	1	ZAN_MOUSE	O88799 mus musculu
45	28	73.7	78	1	VXIS_BPM2	O22002 mycobacteri

ALIGNMENTS

RESULT 1	LV6A_HUMAN	STANDARD;	PRT;	112 AA.
ID	LV6A_HUMAN			
AC	P01721;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ig lambda chain V-VI region AR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE (AMYLLOID PROTEIN AR).			
RX	MEDLINE=82091000; PubMed=6797401;			
RA	Sletten K., Natvig J.B., Husby G., Juul J.;			
RT	"The complete amino acid sequence of a prototype			
RL	immunoglobulin-lambda light-chain-type amyloid-fibril protein AR.";			
RT	Biochem. J. 195:561-572(1981).			
CC	-I- MISCELLANEOUS: ABOUT HALF OF THE LAMBDA CHAIN C REGION IS MISSING			
CC	FROM THIS PROTEIN.			
CC	-I- MISCELLANEOUS: THIS PROTEIN WAS ISOLATED FROM THE SPLEEN OF A			
CC	PATIENT WITH AMYLOIDOSIS.			
CC	-I- SIMILARITY: Contains 1 immunoglobulin-like domain.			
DR	PIR; A01987; L6H0AR.			
DR	HSP; P01709; 2MCG.			
DR	GO; GO:0005576; C:extracellular; NAS.			
DR	GO; GO:0003823; F:antigen binding activity; NAS.			
DR	GO; GO:0006955; P:immune response; NAS.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003596; Ig_V.			
DR	Pfam; PF00047; Ig_1.			
DR	SMART; SM00406; IGV; 1.			
DR	PROSITE; PS50835; IG_LIKE; 1.			
KW	Immunoglobulin V region; Amyloid.			
FT	DOMAIN 1 107			
FT	NON TER 112			
FT	SEQUENCE 112 AA; 11918 MW; 570BCD9A368EF1FE CRC64;			
SO	SEQUENCE			
Query Match	92.1%;			
Best Local Similarity	85.7%;			
Matches	6; Conservative			
Matches	1; Mismatches			
Matches	0; Indels			
Matches	0; Gaps			
Matches	0;			
QY	1 EDNORPS 7			
DB	51 DDNORPS 57			
RESULT 2	LV1F_HUMAN	STANDARD;	PRT;	109 AA.
ID	LV1F_HUMAN			
AC	P04208;			
DT	20-MAR-1987 (Rel. 04, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			

```

DE Ig lambda chain V-I region WAW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=83221661; PubMed=6407018;
RA Takahashi Y., Takahashi N., Teraert D., Putnam F.W.;
RT "Complete covalent structure of a human immunoglobulin D: sequence of
RT the lambda light chain."
RL Proc. Natl. Acad. Sci. U.S.A. 80:3686-3690(1983).
DR HSSP; P01703; 7FAB.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_V.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
DR Immunoglobulin V region.
FT DOMAIN 1 97 V SEGMENT.
FT DISULFID 98 109 J SEGMENT.
FT NON_TER 22 89 BY SIMILARITY.
OX NCBI_TaxID=9606;
SQ SEQUENCE 109 AA; 11725 MW; B17785F6A8DF9BAC CRC64;

Query Match 89.5%; Score 34; DB 1; Length 109;
Best Local Similarity 85.7%; Pred. No. 1.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 EDNORPS 57

RESULT 3
LV6E HUMAN STANDARD; PRT; 111 AA.
ID LV6E HUMAN
AC P06319;
DT 01-JAN-1988 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 01, Last sequence update)
DE Ig lambda chain V-I region WAW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=69060892; PubMed=4177823;
RA Langer B., Steinmetz-Kayne M., Hilschmann N.;
RT "The complete amino acid sequence of Bence Jones protein New (lambda-
RT type). Subgroups in the variable part of immunoglobulin L-chains of
RT the lambda-type."
RL Hoppe-Sejler S.Z. Physiol. Chem. 349:945-951(1968).
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR HSSP; P01703; 7FAB.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_V.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
IMMUNOGLOBULIN V region; Bence-Jones protein;

```

```

KW Pyroliidone carboxylic acid.
FT DOMAIN 1 105 IG-LIKE.
FT MOD_RES 1 1 PYROLIIDONE CARBOXYLIC ACID.
FT DISULFID 22 89 BY SIMILARITY.
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 11453 MW; AAECBCA3C49F2AD3 CRC64;

Query Match 89.5%; Score 34; DB 1; Length 111;
Best Local Similarity 85.7%; Pred. No. 1.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 EDNORPS 57

RESULT 4
LV6E HUMAN STANDARD; PRT; 131 AA.
ID LV6E HUMAN
AC P06319;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1998 (Rel. 38, Last annotation update)
DE Ig lambda chain V-VI region EB4 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85215660; PubMed=3923440;
RA Anderson M.L.M., Brown L., McKenzie E., Kellow J.E., Young B.D.;
RT "Cloning and sequence analysis of an Ig lambda light chain mRNA
RT expressed in the Burkitt's lymphoma cell line EB4."
RL Nucleic Acids Res. 13:2931-2941(1985).
DR HSSP; P01709; 2MCG.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_V.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
IMMUNOGLOBULIN V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 131 IG LAMBDA CHAIN V-VI REGION EB4.
FT DOMAIN 20 41 FRAMEWORK-1.
FT DOMAIN 42 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 110 FRAMEWORK-3.
FT DOMAIN 111 118 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 41 110 FRAMEWORK-4.
FT NON_TER 131 131 BY SIMILARITY.
SQ SEQUENCE 131 AA; 14147 MW; 02A9179C8C05C2CD CRC64;

Query Match 89.5%; Score 34; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 6
DB 70 EDNORPS 75

RESULT 5
LV6D HUMAN STANDARD; PRT; 111 AA.
ID LV6D HUMAN
AC P06318;

```

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig lambda chain V-VI region MLT.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI_Taxid=9606;

OX [1]

RP SEQUENCE.

RX MEDLINE=86122667; PubMed=4089539;

RA Dwulet F.E., Strako K., Benson M.D.;

RT "Amino acid sequence of a lambda VI primary (AL) amyloid protein (MLT)".

RL Scand. J. Immunol. 22:653-660(1985).

DR PIR: A01989; L6HUT.

DR HSSP: P01709; 2MCG.

DR GO; GO:0005576; C:extracellular; NAS.

DR GO; GO:0003823; F:antigen binding activity; NAS.

DR GO; GO:0006955; P:immune response; NAS.

DR InterPro: IPR007110; Ig-like.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_V.

DR Pfam: PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PS50835; IG_LIKE; 1.

KW Immunoglobulin V region.

FT DOMAIN 1 22 FRAMEWORK-1.

FT DOMAIN 23 35 FRAMEWORK-2.

FT DOMAIN 36 50 FRAMEWORK-3.

FT DOMAIN 51 57 FRAMEWORK-4.

FT DOMAIN 58 91 COMPLEMENTARITY-DETERMINING-3.

FT DOMAIN 92 101 FRAMEWORK-4.

FT DISULFID 102 111 BY SIMILARITY.

FT NON TER 111 111

SQ SEQUENCE 111 AA; 11966 MW; 0C88B2FE37BC24F CRC64;

Query Match 86.8%; Score 33; DB 1; Length 111;

Best Local Similarity 85.7%; Pred. No. 2;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7

DB 51 ENNORPS 57

RESULT 6

PEA3 BRARE STANDARD; PRT; 494 AA.

ID PEA3 BRARE

AC 09P01: 057586;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE ETS-domain transcription factor PEA3.

GN PEA3.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OC NCBI_Taxid=7955;

OX [1]

RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, DNA-BINDING, AND REGULATION OF ACTIVITY.

RX TISSUE=Embryo;

RA Brown L.A., Amores A., Schilling T.F., Jowett T., Baert J.-L., de Launait Y., Sharrocks A.D.;

RT "Molecular characterisation of the zebrafish PEA3 ETS-domain transcription factor";

RL Oncogene 17:93-104(1998).

RN [2]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RC TISSUE=Gastrula;

RX MEDLINE=20005800; PubMed=10534622;

RA Muenchberg S.R., Ober E.A., Steinbeiser H.;

RT "Expression of the Ets transcription factors erm and peaz in early zebrafish development";

RL Mech. Dev. 88:233-236(1999).

RN [3]

RP TISSUE SPECIFICITY, AND INDUCTION BY FGF8.

RX MEDLINE=21305906; PubMed=11413000;

RA Roehl H., Nusslein-Volhard C.;

RT "Zebrafish peaz and erm are general targets of FGF8 signaling";

RL Curr. Biol. 11:503-507(2001).

RN [4]

RP TISSUE SPECIFICITY, AND INDUCTION BY FGF3 AND FGF8.

RX MEDLINE=21412239; PubMed=11520667;

RA Raible F., Brand M.;

RT "Tight transcriptional control of the ETS domain factors Erm and Peaz by Fgf signaling during early zebrafish development";

RL Mech. Dev. 107:105-117(2001).

CC -1- FUNCTION: Binds to the (5'-CCGGA/T-3') motif. May control the acquisition of specific cell fates at an early stage during development of the somites and nervous system. May mediate the cellular effects of the fibroblast growth factors on embryogenesis.

CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).

CC -1- TISSUE SPECIFICITY: In the embryo, expressed ubiquitously until the late blastula stage, in the marginal zone of gastrula stages, in the presumptive forebrain and hindbrain and in the trunk region of early somite stages. In later stages, also expressed in Rohon-Beard neurons, epipyls, lateral line placodes, pectoral fin buds, developing lens and heart.

CC -1- DEVELOPMENTAL STAGE: First detected in the embryo after 4.3 hours.

CC -1- INDUCTION: By the fibroblast growth factors FGF3 and FGF8.

CC -1- PTM: Phosphorylated (Probable).

CC -1- MISCELLANEOUS: Transcriptional activation activity is enhanced by Raf1, ERK and PKA.

CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AJ003200; CA05980.1; -.

DR EMBL; AF168008; NAD50434.1; -.

DR HSSP; P14921; 2STT.

DR ZFIN; ZDB-GENE-990415-71; peaz.

DR GO; GO:0005634; C:nucleus; ISS.

DR GO; GO:0015653; F:transcriptional activator activity; IDA.

DR GO; GO:0009790; P:embryonic development; NAS.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.

DR InterPro: IPR000418; Ets.

DR InterPro: IPR006715; HSF_PEA3_N.

DR InterPro: IPR002341; HSF_ETS.

DR Pfam; PF00178; Ets; 1.

DR Pfam; PF04621; ETS_PEA3_N; 1.

DR PRINTS; PR00454; ETSDOMAIN.

DR SMART; SM00413; ETS; 1.

DR PROSITE; PS00345; ETS_DOMAIN_1; 1.

DR PROSITE; PS00346; ETS_DOMAIN_2; 1.

DR PROSITE; PS0061; ETS_DOMAIN_3; 1.

KW DNA-binding; Nuclear protein; Activator; Transcription regulation; Phosphorylation.

FT DOMAIN 47 78 ASP/GLU-RICH (ACIDIC).

FT DNA BIND 350 430 ETS_DOMAIN.

FT CONFLICT 148 148 A -> G (IN REF. 2).

FT CONFLICT 177 177 S -> A (IN REF. 2).

FT CONFLICT 217 218 ID -> MH (IN REF. 2).

FT CONFLICT 224 226 MISSING (IN REF. 2).

FT CONFLICT 392 393 IE -> MQ (IN REF. 2).
 SQ SEQUENCE 494 AA; 55621 MW; 8149253EA25C362 CRC64;
 Query Match
 Best Local Similarity 86.8%; Score 33; DB 1; Length 494;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 DNORPS 7
 Db 443 DNORPS 448
 RESULT 7
 LVEC HUMAN
 ID LIV6C HUMAN STANDARD; PRT; 111 AA.
 AC P06317;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig lambda chain V-VI region SUT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RA Solomon A., Kyle R.A., Frangione B.;
 RT "Light chain variable region subgroups of monoclonal immunoglobulins
 in amyloidosis AL.";
 RL (in) Glenner G.G., Osseman E.F., Benditt E.P., Calkins E.,
 RL Cohen A.S., Zucker-Franklin D. (eds.);
 RL Amyloidosis, pp.449-462, Plenum Press, New York (1986).
 DR PIR; A01988; LGHUS.
 DR PDB; 1C0D; 06-MAR-00.
 DR InterPro; IPR007110; IG_1like.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_V.
 DR Pfam; PF00047; IG_1.
 DR SMART; SM00406; IG_1.
 DR PROSITE; PSS0835; IG_LIKE_1.
 KW Immunoglobulin V region; 3D-structure.
 FT DOMAIN 1 22 FRAMEWORK-1.
 FT DOMAIN 23 35 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 36 50 FRAMEWORK-2.
 FT DOMAIN 51 57 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 58 91 FRAMEWORK-3.
 FT DOMAIN 92 100 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 101 111 FRAMEWORK-4.
 FT DISULFID 22 91 BY SIMILARITY.
 FT NON TER 111 111
 SQ SEQUENCE 111 AA; 12347 MW; 0941DD547D983598 CRC64;
 Query Match
 Best Local Similarity 84.2%; Score 32; DB 1; Length 111;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 EDNORPS 7
 Db 51 EDNORPS 57
 RESULT 8
 YS41_CABEL
 ID YS41_CABEL STANDARD; PRT; 460 AA.
 AC Q09370;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 52.5 kDa protein ZK177.1 in chromosome II.
 GN ZK177.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Anderson K.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U21321; AAC0044.1; -
 DR PIR; T27770; T27770.
 DR WormBep; ZK177.1; CE02089.
 KW Hypothetical protein.
 SQ SEQUENCE 460 AA; 52498 MW; F5E295BCE6CC6F58 CRC64;
 Query Match
 Best Local Similarity 84.2%; Score 32; DB 1; Length 460;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EDNORPS 7
 Db 153 QDNORPA 159
 RESULT 9
 GREB CAMJUE
 ID GREB CAMJUE STANDARD; PRT; 161 AA.
 AC Q9P1K9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcription elongation factor grea (Transcript cleavage factor
 DE grea).
 GN GREB OR C10287C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 CC -1- FUNCTION: Necessary for efficient RNA polymerase transcription
 CC elongation past template-encoded arresting sites. The arresting
 CC sites in DNA have the property of trapping a certain fraction of
 CC elongating RNA polymerases that pass through, resulting in locked
 CC ternary complexes. Cleavage of the nascent transcript by cleavage
 CC factors such as grea or greb allows the resumption of elongation
 CC from the new 3' terminus. GreA releases sequences of 2 to 3
 CC nucleotides (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE GREB/GREB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).


```

CC -----
DR EMBL: AL139074; CAB72754.1; -.
DR PIR: E81447; E81447.
DR HSSP: P21346; 1GRJ.
DR HAMAP: MF_00105; -. 1
DR InterPro: IPR006359; GreA.
DR InterPro: IPR001437; GreA_GreB.
DR Pfam: PF01272; GreA_GreB.1.
DR Pfam: PF03449; GreA_GreB_N.1.
DR ProDom: PD004918; GreA_GreB.1.
DR TIGRfam: TIGR01462; greA.1.
DR PROSITE: PS00829; GREAB_1.1.
DR PROSITE: PS00830; GREAB_2.1.
DR Transcription regulation: DNA-binding; Coiled coil; Complete proteome.
FT DOMAIN 43 74 COILED COIL (POTENTIAL).
SQ SEQUENCE 161 AA; 18010 MW; 132C7DEC962A339 CRC64;

Query Match
Best Local Similarity 71.4%; Score 31; DB 1; Length 161;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 22 KDNORPA 28

RESULT 10
LVSA_HUMAN STANDARD; PRT; 108 AA.
ID P01719;
AC P01719;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig lambda chain V-V region DEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE.
RX MEDLINE=75112179; PubMed=4452363;
RA Bullitt M.;
RT "A new subgroup of human L-chains of the lambda-type. Primary
structure of Bence-Jones protein DEL.",
RL Eur. J. Biochem. 50:48-69(1974)

CC -1- MISCELLANEOUS: THIS IS THE FIRST SEQUENCED V REGION OF LAMBDA
CHAIN SUBGROUP V.
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR HSSP: P80748; 2LQI.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0003823; F:antigen binding activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam: PF00047; Ig.1.
DR SMART: SM00406; IGV.1.
DR PROSITE: PS00835; IG_LIKE.1.
DR Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 97 IG-LIKE.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11342 MW; B8B8BD9C09C9B451 CRC64;

Query Match
Best Local Similarity 71.4%; Score 30; DB 1; Length 108;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 48 EDNDRPA 54

```

```

RESULT 11
YIE2_HSVB4 STANDARD; PRT; 332 AA.
ID YIE2_HSVB4
AC 0024E4;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in IE2 5' region (Fragment).
OS Bovine herpesvirus type 4 (isolate DN-599).
OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=10355;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=93124571; PubMed=8380465;
RA van Santen V.L.;
RT "Characterization of a bovine herpesvirus 4 immediate-early RNA
encoding a homolog of the Epstein-Barr virus R transactivator.",
RL J. Virol. 67:773-784(1993).
CC -1- SIMILARITY: TO EBV BRRF2.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L01099; AAA46056.1; -.
DR PIR: A45710; A45710.
DR Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 332 AA; 37951 MW; E0B6FAP287CC5315 CRC64;

Query Match
Best Local Similarity 83.3%; Score 30; DB 1; Length 332;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORP 6
DB 11 EDNORP 16

RESULT 12
YLII_ECOLI STANDARD; PRT; 371 AA.
ID YLII_ECOLI
AC P75804; Q9R7R6;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yllI precursor.
GN YLII OR B0837.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=X12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III; Bloch C.A.; Perna N.T.; Burland V.;
RA Riley M.; Colado-Vides J.; Glasner J.D.; Rode C.K.; Mayhew G.F.;
RA Gregor J.; Davis N.W.; Kirkpatrick H.A.; Goeden W.A.; Rose D.C.;
RA Mau B.; Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.",
RL Science 277:1245-1247(1997).
CC -----
RP SEQUENCE FROM N.A.
RX STRAIN=X12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,

```

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Maeda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tegtani H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiuchi T.,
 RT "A 718-bp DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12,728,0 min region on the linkage map.",
 RL DNA Res. 3:137-155(1996).
 CC -1- SIMILARITY: STRONG, TO SYNECHOCYSTIS PCC 6803 SLR1608 AND SOME, TO
 CC A.CALCOACETICUS GLUCOSE DEHYDROGENASE-B (GDH).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AE000186; AAC73924.1; -.
 CC DR EMBL: D90721; BAA35532.1; -.
 CC DR EMBL: D90722; BAA35540.1; -.
 CC DR PIR: E64821; E64821.
 CC DR EcoGene: EG13480; Y111.
 CC DR Hypothetical protein; Signal; Complete proteome.
 CC KM SIGNAL 1 20 POTENTIAL.
 CC FT CHAIN 21 371 HYPOTHELTICAL PROTEIN Y111.
 CC SQ SEQUENCE 371 AA; 41054 MW; 99DB08FA302F50B9 CRC64;
 CC
 CC Query Match 78.9%; Score 30; DB 1; Length 371;
 CC Best Local Similarity 71.4%; Pred. No. 36;
 CC Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 EDNORPS 7
 CC ID 166 ENNORPT 172
 CC
 CC RESULT 13
 CC ID ETV4_HUMAN STANDARD; PRT; 551 AA.
 CC AC P43268;
 CC DT 01-NOV-1995 (Rel. 32, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE Adenovirus E1A enhancer binding protein (E1A-F) (ETS translocation
 CC variant 4) (Fragment).
 CC GN ETV4 OR E1AF.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=95293380; PubMed=7774926;
 CC RA Friedman L.S., Ostermeyer E.A., Lynch E.D., Szabo C.I., Meza J.E.,
 RA Anderson L.A., Dowd P., Lee M.K., Rowell S.E., Ellison J.,
 RT Boyd J., King M.-C.,
 RT "2 genes from chromosome 17q21: cloning, sequencing, and
 RT characterization of mutations in breast cancer families and tumors.",
 RL Genomics 25:256-263 (1995).
 CC [2]
 CC RN SEQUENCE OF 90-551 FROM N.A.
 CC RX MEDLINE=93161246; PubMed=8441666;
 CC RA Higashino F., Yoshida K., Fujinaga K., Kamio K., Fujinaga K.,
 RT "Isolation of a cDNA encoding the adenovirus E1A enhancer binding
 RT protein: a new human member of the ets oncogene family.",
 RL Nucleic Acids Res. 21:547-553(1993).
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO THE ENHANCER OF
 CC THE ADENOVIRUS E1A GENE; THE CORE-BINDING SEQUENCE IS
 CC 5' [AC]GGA[AT]GT-3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: U18018; AAA95991.1; -.
 CC DR EMBL: D12765; BAA02234.1; -.
 CC DR HSSP: P28324; 1BC8.
 CC DR TRANSFAC: T00685; -.
 CC DR Genew; HGNC:3493; ETV4.
 CC DR MIM: 600711; -.
 CC DR GO: GO:0005634; C:nucleus; NAS.
 CC DR GO: GO:0015633; F:transcriptional activator activity; NAS.
 CC DR GO: GO:000355; F:regulation of transcription, DNA-dependent; NAS.
 CC DR InterPro: IPR000418; Ets.
 CC DR InterPro: IPR006715; Ets_PEA3_N.
 CC DR InterPro: IPR002341; HSF_Ets.
 CC DR Pfam: PF00178; Ets_1.
 CC DR Pfam: PF04621; Ets_PEA3_N; 1.
 CC DR PRINTS: PR00454; ETSDOMAIN.
 CC DR SMART: SM00413; ETS; 1.
 CC DR PROSITE: PS00345; ETS DOMAIN 1; 1.
 CC DR PROSITE: PS00346; ETS DOMAIN 2; 1.
 CC DR PROSITE: PS50661; ETS DOMAIN 3; 1.
 CC KM DNA-binding; Nuclear protein; Activator; Transcription regulation.
 CC FT NON_TER 1 1
 CC FT DOMAIN 116 142 ASP/GLU-RICH (ACIDIC).
 CC FT DOMAIN 215 311 GLN-RICH.
 CC FT DNA_BIND 408 488 ETS-DOMAIN.
 CC FT CONFLICT 91 94 GNGS -> EMSD (IN REF. 2).
 CC SQ SEQUENCE 551 AA; 60570 MW; 9D649C8A1FBDBDB CRC64;
 CC
 CC Query Match 78.9%; Score 30; DB 1; Length 551;
 CC Best Local Similarity 83.3%; Pred. No. 56;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 DNORPS 7
 CC ID 501 DNORPA 506
 CC
 CC RESULT 14
 CC ID ETV4_MOUSE STANDARD; PRT; 555 AA.
 CC AC P28332;
 CC DT 01-DEC-1992 (Rel. 24, Created)
 CC DT 01-DEC-1992 (Rel. 24, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Polyomavirus enhancer activator 3 (PEA3 protein) (ETS translocation
 CC variant 4).
 CC GN ETV4 OR PEA3 OR PEA-3.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=92192459; PubMed=1547944;
 CC RA Xin J.-H., Cowie A., Lachance P., Hassell J.A.,
 RT "Molecular cloning and characterization of PEA3, a new member of the
 RT Ets oncogene family that is differentially expressed in mouse
 RT embryonic cells.",
 RL Genes Dev. 6:481-496(1992).
 CC -1- FUNCTION: BINDS TO THE PEA3 MOTIF (5'-AGGAG-3'). MAY PLAY A
 CC REGULATORY ROLE DURING EMBRYOGENESIS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: EPIDIDYMIS AND THE BRAIN.
 CC -1- PTM: PHOSPHORYLATED.
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X63190; CAA44872.1; -
CC PIR; S24061; S24061.
CC HSSP; P28324; 1BC8.
CC TRASPAC; T00684; -
CC MGD; GGI:99423; ECV4.
CC GO; GO:0008045; P:motor axon guidance; IMP.
CC InterPro; IPR000418; Ets_PEA3_N.
CC InterPro; IPR006715; Ets_PEA3_N.
CC InterPro; IPR002341; HSF_ET5.
CC Pfam; PF00178; Ets; 1.
CC PRINTS; PR004621; Ets_PEA3_N; 1.
CC SMART; SM00413; ETS; 1.
CC PROSITE; PS00345; ETS_DOMAIN_1; 1.
CC PROSITE; PS00346; ETS_DOMAIN_2; 1.
CC PROSITE; PS50061; ETS_DOMAIN_3; 1.
CC DNA-binding; Nuclear protein; Activator; Transcription regulation;
CC Phosphorylation.
CC DOMAIN 124 150 ASP/GU-RICH (ACIDIC).
CC FT BIND 217 315 GIN-RICH.
CC FT DNA_BIND 412 492 ETS-DOMAIN.
CC SQ SEQUENCE 555 AA; 60846 MW; 278076B4A8A4D4B2 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 555;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
Db 505 DNORPS 510

RESULT 15
ID ITK_HUMAN STANDARD; PRT; 620 AA.
AC 00861;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase ITK/TSK (EC 2.7.1.112) (T-cell-specific
DE kinase) (Tyrosine-protein kinase Ltk) (Kinase EMT).
GN ITK OR LTK OR EMT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93279365; PubMed=8504851;
RA Tanaka N., Asao H., Ohtani K., Nakamura M., Sugamura K.;
RT "A novel human tyrosine kinase gene inducible in T cells by
RT Interleukin 2,"
RL FEBS Lett. 324:1-5(1993).
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
MEDLINE=93372354; PubMed=8364206;
RA Gibson S., Leung B., Squire J.A., Hill M., Arima N., Goss P.,
RA Hogg D., Mills G.B.;
RT "Identification, cloning, and characterization of a novel human
RT T-cell-specific tyrosine kinase located at the hematopoietic complex
RT on chromosome 5q,"
RL Blood 82:1561-1572(1993).
CC -1- FUNCTION: PLAYS A ROLE IN T CELL PROLIFERATION AND

CC DIFFERENTIATION.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: LOCALIZES TO CELL SURFACE RECEPTORS IN THE
CC PLASMA MEMBRANE AFTER STIMULATION WITH RESPECTIVE RECEPTORS (TCR,
CC CD28, CD2) IN T-CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: T-CELL LINES AND NATURAL KILLER CELL LINES.
CC -1- INDUCTION: By Interleukin-2.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. TEC
CC SUPERFAMILY.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -1- SIMILARITY: Contains 1 PH domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; D13720; BAA02873.1; -
CC EMBL; L10717; AAA36748.1; -
CC EMBL; S65186; AAB28072.2; -
CC PIR; S33253; S33253.
CC HSSP; O03526; 1AM7.
CC GeneW; HGNC:6171; ITK.
CC MIM; 166973; -
CC GO; GO:0004715; F:non-membrane spanning protein tyrosine kina. . .; TAS.
CC GO; GO:0006968; P:cellular defense response; TAS.
CC InterPro; IPR001562; BTK.
CC InterPro; IPR001849; PH.
CC InterPro; IPR000719; PTK.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00779; BTK; 1.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00402; TECHDOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC PRODOM; PD000001; PTK_kinase; 1.
CC PRODOM; PD000093; SH2; 1.
CC PRODOM; PD000066; SH3; 1.
CC SMART; SM00107; BTK; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 1.
CC SMART; SM00219; TYRK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50002; SH2; 1.
CC PROSITE; PS50002; SH3; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC TRANSFERASE; Tyrosine-protein kinase; Phosphorylation;
CC ATP-binding; SH2 domain; SH3 domain.
CC FT DOMAIN 4 111 PH.
CC FT DOMAIN 171 231 SH3.
CC FT DOMAIN 239 338 SH2.
CC FT DOMAIN 363 615 PROTEIN_KINASE.
CC FT NP_BIND 369 377 ATP (BY SIMILARITY).
CC FT BINDING 391 391 ATP (BY SIMILARITY).
CC FT ACT_SITE 482 482 BY SIMILARITY.
CC FT MOD_RES 512 512 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC FT CONFLICT 331 331 V -> W (IN REF. 2; AAB28072).
CC SQ SEQUENCE 620 AA; 71831 MW; DAE396BD2309319D CRC64;

Query Match 78.9%; Score 30; DB 1; Length 620;

Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNQR 6
|||:
Db 162 EDNR 167

Search completed: January 12, 2004, 06:59:45
Job time : 5.82812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:04 ; Search time 14.2168 Seconds
(without alignments)
127.041 Million cell updates/sec

Title: US-09-829-495-65

Perfect score: 38
Sequence: 1 EDNORPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPRENBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	112	4 Q96CUD1	Q96JDI homo sapien
2	38	100.0	116	4 Q96CUD0	Q96JDI homo sapien
3	34	89.5	248	4 Q8WMI8	Q8WMI8 homo sapien
4	34	89.5	266	2 Q49849	Q49849 mycobacteri
5	34	89.5	704	10 Q80896	Q80896 arachidopis
6	33	86.8	112	4 Q96CUD2	Q96JDI homo sapien
7	33	86.8	222	2 Q93D14	Q93D14 acinetobact
8	32	86.8	478	5 Q43955	Q43955 leishmania
9	32	84.2	471	5 Q09534	Q09534 caenorhabdi
10	32	84.2	603	16 Q82722	Q82722 salmonella
11	32	84.2	691	16 Q82PF2	Q82PF2 salmonella
12	31	81.6	193	16 Q8B075	Q8B075 oceanobacti
13	31	81.6	221	13 Q90X42	Q90X42 brachydanio
14	31	81.6	233	4 Q8TBC9	Q8TBC9 homo sapien
15	31	81.6	250	16 Q8DV81	Q8DV81 streptococc
16	31	81.6	395	16 Q8G873	Q8G873 bifidobacte

17	31	81.6	450	2 Q8VQV1	Q8VQV1 rhodococcus
18	31	81.6	1097	5 Q9V4C9	Q9V4C9 drosophila
19	30	78.9	85	2 Q919B7	Q919B7 chlamydia t
20	30	78.9	150	16 Q8EXI6	Q8EXI6 leptospira
21	30	78.9	158	16 Q97060	Q97060 streptococc
22	30	78.9	158	16 Q8DPD4	Q8DPD4 streptococc
23	30	78.9	190	2 Q8GHN3	Q8GHN3 pseudomonas
24	30	78.9	190	4 Q96T00	Q96T00 homo sapien
25	30	78.9	202	12 Q98VM3	Q98VM3 rabies viru
26	30	78.9	232	10 Q22922	Q22922 arachidopis
27	30	78.9	265	13 Q98SV0	Q98SV0 brachydanio
28	30	78.9	336	13 Q9YH55	Q9YH55 gallus galli
29	30	78.9	335	16 Q9F3F6	Q9F3F6 streptomyc
30	30	78.9	366	2 Q9YMR9	Q9YMR9 lymantria d
31	30	78.9	366	2 Q9YMR9	Q9YMR9 neisseria m
32	30	78.9	371	16 Q8X6U6	Q8X6U6 escherichia
33	30	78.9	371	16 Q8CW87	Q8CW87 escherichia
34	30	78.9	390	9 Q9FZK6	Q9FZK6 mycoplasma
35	30	78.9	400	5 Q963B5	Q963B5 loligo peal
36	30	78.9	409	10 Q9SGC6	Q9SGC6 arachidopis
37	30	78.9	415	11 Q8CAF0	Q8CAF0 mus musculu
38	30	78.9	420	10 Q9SJP6	Q9SJP6 arachidopis
39	30	78.9	484	4 Q96AM9	Q96AM9 homo sapien
40	30	78.9	514	12 Q99CZ2	Q99CZ2 bovine herp
41	30	78.9	522	2 Q85210	Q85210 vibrio chol
42	30	78.9	522	16 Q9K394	Q9K394 vibrio chol
43	30	78.9	555	12 Q91HJ5	Q91HJ5 avian param
44	30	78.9	561	10 Q8GW98	Q8GW98 arachidopis
45	30	78.9	664	10 Q64589	Q64589 arachidopis

ALIGNMENTS

RESULT 1
ID Q96CUD1 PRELIMINARY; PRT; 112 AA.
AC Q96CUD1;
DT 01-DEC-2001 (TRMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRMBLrel. 23, Last annotation update)
DE Amyloid lambda 6 light chain variable region PIP (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Perlefin V., Casarini S., Colla Vignarelli M., Merlini G.;
RT "Amyloid lambda 6 light chain variable region PIP";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF267874; AAK58586.1; -
DR InterPro; IPR007110; IG-1-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; igf_1.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12047 MW; 0D385AC23567B9F CRC64;

Query Match 100.0%; Score 38; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 0.73; Indels 0;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 1 EDNORPS 7
|||||
Db 51 EDNORPS 57

RESULT 2

Q96JD0
ID Q96JD0 PRELIMINARY; PRT; 116 AA.
AC Q96JD0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Amyloid lambda 6 light chain variable region SAR (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Perfecti V., Casarini S., Colli Vignarelli M., Merlini G.;
RT "Amyloid lambda 6 light chain variable region SAR."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF267875; AAK58587.1; -
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig_1.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT 116 116
SQ SEQUENCE 116 AA; 12294 MW; F7B0E9F49FAE369E CRC64;
Query Match 100.0%; Score 38; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORPS 7
Db 51 EDNORPS 57
RESULT 3
08MW18 PRELIMINARY; PRT; 248 AA.
ID Q8MW18
AC Q8MW18;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Similar to RIKEN cDNA 1700020L24 gene.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021727; AAH21727.1; -
SQ SEQUENCE 248 AA; 27236 MW; A988860B81AA530D CRC64;
Query Match 89.5%; Score 34; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORP 6
Db 37 EDNORP 42
RESULT 4
049849 PRELIMINARY; PRT; 266 AA.
ID Q49849
AC Q49849;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE B2235_F2_77.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U00019; AAA17283.1; -
SQ SEQUENCE 266 AA; 29602 MW; 02C50A1B4E36EC82 CRC64;
Query Match 89.5%; Score 34; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORP 6
Db 103 EDNORP 108
RESULT 5
080896 PRELIMINARY; PRT; 704 AA.
ID Q80896
AC Q80896;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE At2g32590 protein.
GN AT2G32590.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bentio M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanKen S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams W.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana."
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004681; AAC25941.1; -
SQ SEQUENCE 704 AA; 79017 MW; 69C8E8FD5CC040FA CRC64;
Query Match 89.5%; Score 34; DB 10; Length 704;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORPS 7
Db 330 EDNORPS 336
RESULT 6
Q96JD2 PRELIMINARY; PRT; 112 AA.
ID Q96JD2

```
AC 096D2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Amyloid lambda 6 light chain variable region NEG (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow;
RA Perfect V., Casarini S., Colli Vignarelli M., Merlini G.;
RT "Amyloid lambda 6 light chain variable region NEG."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF267873; AAK58585.1; -.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
FT NON_TER 1
FT 112 112
SQ SEQUENCE 112 AA; 11908 MW; 080B4B37E2360B06 CRC64;

Query Match
Best Local Similarity 86.8%; Score 33; DB 4; Length 112;
Pred. No. 10;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNQRPS 7
DB 51 ENNRQPS 57

RESULT 7
ID 093D14 PRELIMINARY; PRT; 222 AA.
AC 093D14;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 20, last annotation update)
DE O-methyltransferase-like protein.
OS Acinetobacter sp. ADP1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ADP1;
RX MEDLINE=21555098; PubMed=11698371;
RA Young D.M., Ornston L.N.;
RT "Functions of the Mismatch Repair Gene mutS from Acinetobacter sp.
RT Strain ADP1."
RL J. Bacteriol. 183:6822-6831(2001).
DR EMBL: AF400582; AAK92498.1; -.
DR InterPro: IPR002935; Methyltransf_3.
DR InterPro: IPR000051; SAM_bind.
DR Pfam: PF01596; Methyltransf_3; 1.
DR Methyltransferase; Transferase.
SQ SEQUENCE 222 AA; 24730 MW; E8C020F41AD3CF67 CRC64;

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 222;
Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DNQRPS 7
DB 177 DNQRPS 182

RESULT 8
O43955
```

```
ID 043955 PRELIMINARY; PRT; 478 AA.
AC 043955;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, last annotation update)
DE NupML.
GN TOR.
OS Leishmania mexicana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5665;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=LV 78;
RX MEDLINE=98135664; PubMed=9476797;
RA Dekte S.;
RT "Identification of a transcription factor like protein at the TOR
RT locus in Leishmania mexicana amazonensis."
RL Mol. Biochem. Parasitol. 90:505-511(1997).
DR EMBL: AF016581; AAC38851.1; -.
SQ SEQUENCE 478 AA; 52237 MW; 60363D4E0210CF61 CRC64;

Query Match
Best Local Similarity 86.8%; Score 33; DB 5; Length 478;
Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DNQRPS 7
DB 342 DNQRPS 347

RESULT 9
ID 009534 PRELIMINARY; PRT; 471 AA.
AC 009534;
DT 01-JAN-1999 (TReMBLrel. 09, Created)
DT 01-JAN-1999 (TReMBLrel. 09, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE F10B5.3 protein.
GN F10B5.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)
RP SEQUENCE FROM N.A.
RA Sims W.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z48334; CAA88310.1; -.
DR WormBep: F10B5.3; CE01545.
DR InterPro: IPR000210; BTB_POZ.
DR InterPro: IPR007087; ZNF_C2H2.
DR Pfam: PF00096; ZF-C2H2; 3.
DR SMART: SM00225; BTB; 1.
DR SMART: SM00355; ZNF_C2H2; 4.
DR PROSITE: PS50097; BTB; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
DR Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 471 AA; 51751 MW; C9F7B0E3AE882CA CRC64;

Query Match
Best Local Similarity 84.2%; Score 32; DB 5; Length 471;
Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNQRPS 7
DB 3 EDNQRPS 9
```

RESULT 10

082722 PRELIMINARY; PRT; 603 AA.

ID 082722;
AC 082722;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Putative glycogen debranching protein (EC 3.2.1.-).
GN STY1505 OR STY4273.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_Taxid=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mangall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltham T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.,
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AL627270; CAD01764.1; -;
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR004193; Glyco_hydro_13N.
DR Pfam; PF00128; alpha-amyase_1.
DR Pfam; PF02922; isocamylase_N; 1.
DR POSITE; PS00018; EF_HAND; 1.
KM Hydrolase; Glycosidase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 603 AA; 66588 MW; 0D0B16B5D6C6B2 CRC64;

Query Match 84.2%; Score 32; DB 16; Length 603;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 154 EDRORPS 160

RESULT 11

082PF2 PRELIMINARY; PRT; 691 AA.

ID 082PF2;
AC 082PF2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Putative glycosyl hydrolase (EC 3.2.1.-).
GN STM1558.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_Taxid=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SCS1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";

RL Nature 413:852-856(2001).

DR EMBL; AE008768; AL20476.1; -;
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR004193; Glyco_hydro_13N.
DR Pfam; PF00128; alpha-amyase_1.
DR Pfam; PF02922; isocamylase_N; 1.
DR POSITE; PS00018; EF_HAND; 1.
KM Hydrolase; Glycosidase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 691 AA; 78569 MW; 318863462FBA6531 CRC64;

Query Match 84.2%; Score 32; DB 16; Length 691;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 154 EDRORPS 160

RESULT 12

08275 PRELIMINARY; PRT; 193 AA.

ID 08275;
AC 08275;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
DE Transposase for 1866 (divided with OB1837 and OB1838).
GN OB1838.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_Taxid=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004599; BAC13794.1; -;
KM Complete proteome.
SQ SEQUENCE 193 AA; 23201 MW; 6CFA619F8DAEB61C CRC64;

Query Match 81.6%; Score 31; DB 16; Length 193;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 49 EDNGRPS 55

RESULT 13

090X42 PRELIMINARY; PRT; 221 AA.

ID 090X42;
AC 090X42;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE SC:dz150F13.1 (Novel protein) (Fragment).
GN SC:dz150F13.1.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC Corby N.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL591213; CAD10077.1; -;

FT NON TER 1 1
SQ SEQUENCE 221 AA; 25208 MW; F0B3D4CAD91CACE1 CRC64;

Query Match 81.6%; Score 31; DB 13; Length 221;
Best Local Similarity 71.4%; Pred. No. 59;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
Db 141 EENOKPS 147

RESULT 14

O8TBC9 PRELIMINARY; PRT; 233 AA.

ID O8TBC9

AC O8TBC9;

DT 01-JUN-2002 (TRENBLREL. 21, Created)

DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)

DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

NCBI_TaxID=9606;

RP TISSUE=B-cell;

RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC022823; AAH22823.1; -.

DR InterPro; IPR003006; IG_MHC.

DR Pfam; PF00047; IG_2.

DR SMART; SM00406; TCV; 1.

DR PROSITE; PS00290; IG_MHC; 1.

KW Hypothetical protein.

SQ SEQUENCE 233 AA; 24867 MW; 367411BFD6F4DF92 CRC64;

Query Match 81.6%; Score 31; DB 4; Length 233;
Best Local Similarity 71.4%; Pred. No. 63;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
Db 68 KDNORPS 74

RESULT 15

O8DV81

ID O8DV81 PRELIMINARY; PRT; 250 AA.

AC O8DV81;

DT 01-MAR-2003 (TRENBLREL. 23, Created)

DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)

DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)

DE Putative 1-acylglycerol-3-phosphate O-acyltransferase (EC

2.3.1.51).

GN SMU.624.

OC Streptococcus mutans.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

NCBI_TaxID=1309;

RP TISSUE=B-cell;

RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC022823; AAH22823.1; -.

DR InterPro; IPR003006; IG_MHC.

DR Pfam; PF00047; IG_2.

DR SMART; SM00406; TCV; 1.

DR PROSITE; PS00290; IG_MHC; 1.

KW Hypothetical protein.

SQ SEQUENCE 250 AA; 29051 MW; 412A631E431A2C3B CRC64;

Query Match 81.6%; Score 31; DB 16; Length 250;
Best Local Similarity 71.4%; Pred. No. 68;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
Db 201 KDNORPS 207

Search completed: January 12, 2004, 07:02:13
Job time : 19.218 secs

THIS PAGE BLANK (USPTO)

PI	Bunfield SJ, Villalal J, Jandrot-Perrus M, Vainchencker W, Gill DS,
PI	Qian MD, Kingsbury G;
DR	WPI; 2001-080877/09.
XX	
PT	New genes encoding human platelet-expressed collagen receptor,
PT	glycoprotein VI, and its modulators, useful for preventing, treating
PT	and diagnosing hemorrhagic disorders, thrombotic diseases and
XX	immunological disorders -
XX	
PS	Claim 32; Page 102; 227pp; English.
XX	
CC	The present sequence is given in a specification relating to an isolated
CC	nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC	glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
CC	and polypeptides and their modulators, e.g. antisense nucleic acids,
CC	ribozymes and antibodies, are useful for preventing, treating and
CC	diagnosing disorders associated with aberrant expression or activity of
CC	GPVI. These disorders include bleeding disorders
CC	(e.g. thrombocytopenia), blood vessel injury, thrombotic disorders
CC	(e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC	disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC	ischæmia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC	infarction), immunological diseases (e.g. platelet disorder) and
CC	embryonic liver disorders. Preferably they are used to prevent acute
CC	cardiac ischaemia following angioplasty and metastatic cancers,
CC	especially of the colon and liver.
XX	
SQ	Sequence 7 AA;
XX	
Query Match	100.0%; Score 38; DB 22; Length 7;
Best Local Similarity	100.0%; Pred. No.9.3e+05;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps
XX	
QY	1 EDNORPS 7
1 EDNORPS 7	
DB	
ABU11263	
ID	ABU11263 standard; Peptide; 7 AA.
XX	
XX	ABU11263;
XX	
DT	06-FEB-2003 (first entry)
XX	
DE	Human TANGO 268 VLCDRI Peptide #3.
XX	
XX	Human; mouse; variable heavy; VH; antigen; cancer;
KW	complementarily determining region; TANGO 268; glycoprotein VI; GPVI;
KW	TANGO 268; extracellular matrix; collagen; platelet release;
KW	proliferation; migration; embryogenesis; inflammation; thrombosis;
KW	deregulation; thrombocytopenia; antibody; thrombotic disorder;
KW	cerebral vascular disease; stroke; ischaemia; venous thromboembolism;
KW	leg swelling; pain; ulceration; pulmonary embolism; coronary disease;
KW	cardiovascular disease; angina pectoris; myocardial infarction;
KW	coronary revascularisation; atherosclerosis; immunological disorder;
KW	developmental disorder; embryonic disorder; liver disorder;
KW	cerebral vascular disease; venous thromboembolism disease.
XX	
OS	Homo sapiens.
XX	
PN	WO200280968-A1.
XX	
PD	17-OCT-2002.
XX	
XX	09-APR-2002; 2002WO-US11122.
XX	
XX	09-APR-2001; 2001US-0829495.
XX	
XX	(MILL-) MILLENNIUM PHARM INC.
XX	

Pt Bluefield SJ, Villaveal J, Jandrot-Pernus M, Vainchencker W, Gill DS;
Pi Qian DM, Kingebury G;
XX WPJ; 2003-058477/05.

XX Novel substantially purified antibody immunospecifically binding to
PT TANGO 268 antigen, useful for treating bleeding disorders such as
PT thrombocytopenia, stroke, ischemia, pulmonary embolism, atherosclerosis

PS Claim 11, Page 111, 236pp; English.

XX This invention relates to a novel purified antibody comprising a
CC variable heavy (VH) complementarity determining region (CDR1, VH CDR2
CC or VH CDR3; or variable light (VL) CDR1, VL CDR2 or VL CDR3, and
CC immunospecifically binding to a TANGO 268 (also referred as glycoprotein
CC VI (GPVI)) antigen. The antibodies of the invention act to decrease or
CC block TANGO 268 binding to extracellular matrix components, or as a
CC collagen or platelet release and aggregation blocker. The antibodies of
CC the invention are useful for modulating proliferation, migration,
CC morphology, differentiation and/or function of megakaryocytes and
CC platelets, including during development e.g. embryogenesis, modulating
CC leukocyte-platelet and platelet-endothelium interactions in
CC inflammation and/or thrombosis, and modulating platelet aggregation and
CC degranulation. They are also useful for modulating disorders associated
CC with abnormal or aberrant megakaryocyte and/or platelet proliferation,
CC migration, morphology, differentiation and/or function, e.g. bleeding
CC disorders such as thrombocytopaenia. Other diseases which may be
CC modulated by these antibodies are thrombotic disorders, cerebral
CC vascular diseases (e.g. stroke and ischaemia) venous thromboembolism
CC diseases (e.g. diseases involving leg swelling, pain and ulceration,
CC pulmonary embolism, etc); coronary diseases (e.g. cardiovascular
CC diseases including angina pectoris, myocardial infarction, coronary
CC reestenosis, atherosclerosis), etc); immunological disorders,
CC rheumatoid arthritis, embryonic disorders, liver disorders, cerebral
CC developmental disorder, venous thromboembolism disease, coronary diseases,
CC vascular diseases, venous thromboembolism disease, coronary diseases,
CC and metastatic cancers. The antibodies of the invention only causes a
CC transient decrease in platelet counts, platelet aggregation, and/or
CC platelet activation and so have some advantages over prior art
CC methods. The present sequence represents a peptide sequence used to
CC generate the antibodies of the invention.

SQ Sequence 7 AA;

Cy Query Match 100.0%; Score 38; DB 24; Length 7;
Beat Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 1 EDNORPS 7
1 EDNORPS 7

RESULT 3
ABP56510 ID ABP56510 standard; Protein: 98 AA.
XX ABP56510;
AC ACP56510;
XX 20-MAR-2003 (first entry)
DT Human anti-Fc-epsilon-R1 alpha autoantibody light chain VI-22.
DE Autoantibody; Fc-epsilon-R1 receptor alpha-chain; immunosuppressive;
XX allergic disease; urticaria; late phase allergic reaction; malignancy;
KW intrinsic asthma; drug intolerance; food intolerance; immunoglobulin E;
conditional autoimmunity; IgE mediated disease.
XX Homo sapiens.
OS Synthetic.
XX WO200282085-A2.

PD 17-OCT-2002.
 XX
 XX 03-APR-2002; 2002WO-EP03660.
 XX
 XX 04-APR-2001; 2001US-281024P.
 XX
 XX (ZLBB-) ZLB BIOPLASMA AG.
 PA
 XX
 XX Miescher S;
 XX
 XX WPI; 2003-103348/09.
 DR
 XX
 XX Identifying and obtaining inhibitor of a pathological process for
 PT creating e.g. autoimmunity comprises determining if a compound is
 PT capable of modulating the binding of the Fc-epsilon-R1 receptor and an
 PT autoantibody against its alpha-chain -
 XX
 XX Claim 20; Page 22; 29pp; English.
 PS
 XX The present invention describes a method for identifying and obtaining
 CC an inhibitor of a pathological process. The method comprises determining
 CC if a compound is capable of modulating the binding of the Fc-epsilon-R1
 CC receptor, alpha-chain and an autoantibody against its alpha-chain. Also
 CC described: (1) use of the autoantibody against the Fc-epsilon-R1 receptor
 CC alpha-chain for identifying and obtaining an inhibitor of a pathological
 CC process; (2) use of the identified inhibitor for inhibiting activity of
 CC the autoantibody against the Fc-epsilon-R1 receptor alpha-chain; and
 CC (3) a compound identified by the method, which binds but does not
 CC activate the receptor; and (4) a polypeptide capable of specific binding
 CC to the Fc-epsilon-R1 receptor alpha-chain. The method is useful for
 CC obtaining an inhibitor of a pathological process e.g. imbalance between
 CC cell-bound and free IGE e.g. allergic disease (urticaria, late phase
 CC allergic reactions, intrinsic asthma, drug intolerance and food
 CC intolerance), IGE mediated disease or malignancy. The compound is useful
 CC for treating a pathological process, particularly conditional
 CC autoimmunity. The present sequence represents a human recombinant
 CC anti-Fc-epsilon-R1 alpha autoantibody light chain protein sequence from
 CC the present invention.
 CC
 XX
 SQ Sequence 98 AA;
 Query Match 100.0%; Score 38; DB 24; Length 98;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDNORPS 7
 DB 51 EDNORPS 57
 RESULT 4
 AAU02522
 ID AAU02522 standard; Protein; 113 AA.
 XX
 AC AAU02522;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 XX Anti-adipocyte monoclonal antibody light chain, FAT 13.
 DE
 XX Antibody; adipocyte; heavy chain; light chain; obesity; fat;
 KW heart disease; complementarity determining region; CDR.
 XX
 OS Homo sapiens.
 XX
 XX WO200127279-A1.
 PN
 PD 19-APR-2001.
 XX
 PF 11-OCT-2000; 2000WO-GB03900.
 XX
 PR 12-OCT-1999; 99US-0158812.
 XX

PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 XX Edwards BM, Main SH, Vaughan TJ;
 XX
 XX WPI; 2001-282031/29.
 DR N-PSDB; AA03422.
 XX
 XX Panel of specific binding members of antibody molecules which bind to
 PT whole adipocytes is used in the treatment of obesity and obesity
 PT related diseases -
 XX
 XX Claim 1; Page 106; 182pp; English.
 PS
 XX AAU02501-AAU02635, and AAU02641-AAU02748 represent the amino acid
 CC sequences of anti-adipocyte monoclonal antibody heavy chain, light
 CC chain, and heavy chain complementarity determining regions (CDR) of the
 CC invention. The antibodies can be used in the treatment of obesity and
 CC obesity related diseases. The antibodies can be used to deliver drugs or
 CC pro-drugs directly to the fat mass of an obese patient or the antibody
 CC can be used as a therapeutic itself. Antibodies binding specifically to
 CC adipocytes can be used to activate the immune system to destroy the cells
 CC by complement mediated lysis. The antibodies may be labeled with a
 CC detectable label such as radiolabel, fluorescent or chemical group and
 CC used in methods of diagnosis in human subjects e.g. to determine the
 CC presence of adipocyte antigen on the surface of an adipocyte to detect or
 CC determine the presence or level of adipocytes in a cell or tissue sample.
 CC The antibodies can be used as an alternative means of treatment for obese
 CC patients other than undergoing surgery to remove excess fat. Antibodies
 CC for different types of fat deposits can also be produced e.g. intra-
 CC abdominal fat associated with heart disease.
 CC
 XX
 SQ Sequence 113 AA;
 Query Match 100.0%; Score 38; DB 22; Length 113;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDNORPS 7
 DB 53 EDNORPS 59
 RESULT 5
 AAG76011
 ID AAG76011 standard; Protein; 161 AA.
 XX
 AC AAG76011;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 XX Human colon cancer antigen protein SEQ ID NO:6775.
 DE
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 22.
 XX
 OS Homo sapiens.
 XX
 XX WO200122920-A2.
 PN
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CF, Rosen CA;
 PI
 XX WPI; 2001-235357/24.
 DR N-PSDB; AAH35416.
 XX

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11, Page 8234-8235; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing P.
 CC Inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 161 AA;

Query Match 100.0%; Score 38; DB 22; Length 161;
 Best Local Similarity 100.0%; Pred. No. 5.1; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
 |||||
 Db 72 EDNORPS 78

RESULT 6
 ABP45223
 ID ABP45223 standard; Protein; 258 AA.

XX ABP45223;
 AC
 DT 19-AUG-2002 (first entry)
 XX

DE Human BlyS binding scFv SEQ ID 1234.

XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KM tumour necrosis factor; B cell proliferation; B cell differentiation;
 KM immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KM antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KM systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KM common variable immunodeficiency; acquired immunodeficiency syndrome.

XX Homo sapiens.

XX WO200202641-A1.

XX 10-JAN-2002.

XX 15-JUN-2001; 2001WO-US19110.

XX 16-JUN-2000; 2000US-212210P.

XX 17-OCT-2000; 2000US-240816P.

XX 16-MAR-2001; 2001US-276248P.

XX 21-MAR-2001; 2001US-273799P.

XX 25-MAY-2001; 2001US-293499P.

XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTI-BODY TECHNOLOGY.

XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX WPI; 2002-114799/15.

XX Antibodies against B lymphocyte Stimulating polypeptides, useful for
 PT the diagnosis and treatment of cancers and immune disorders -
 XX
 PS Claim 1, Page 1874-1875; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to
 CC B lymphocyte stimulator (BlyS) polypeptides. BlyS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
 CC and so may be used to detect and quantify the presence of BlyS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BlyS. They may also be
 CC administered to treat diseases associated with aberrant BlyS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis, and
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method
 CC of the invention.

XX Sequence 258 AA;

Query Match 100.0%; Score 38; DB 23; Length 258;
 Best Local Similarity 100.0%; Pred. No. 8.5; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
 |||||
 Db 198 EDNORPS 204

RESULT 7
 AAM16592
 ID AAM16592 standard; Peptide; 7 AA.

XX AAM16592;
 AC

DT 30-NOV-1997 (first entry)
 XX

DE Anti-RSV F glycoprotein antibody RSVF2-5 light chain VL6 CDR2.

XX Respiratory syncytial virus; RSV; monoclonal antibody; CDR;
 KM complementarity determining region; pneumonia; bronchiolitis;
 KM diagnosis; therapy; vaccine; RSVF2-5.

XX Homo sapiens.

XX WO9710846-A1.

XX 27-MAR-1997.

XX 18-SEP-1996; 96WO-US14937.

XX 18-SEP-1995; 95US-0003931.

XX (INTR-) INTRACEL CORP.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Chanock RM, Crowe JE, Gilmour PS, Murphy BR, Pilkington GR;
 XX WPI; 1997-202621/18.

XX 1997-202621/18.

XX Composition comprising respiratory syncytial virus antibody - useful
 PT for treatment or prophylaxis of active disease or infection

XX Claim 8; Page 56; 71pp; English.

XX This peptide sequence comprises complementarity determining region
 CC 2 (CDR2) of the light chain VL6 region of a novel neutralising

CC human monoclonal antibody to the respiratory syncytial virus (RSV).
CC The antibody, designated RSVF2-5 (ATCC 6909), selectively binds to
CC an RSV F glycoprotein epitope. DNA encoding RSVF2-5 Fd and light
CC chain (see AAT6556-57) was isolated from a phage library prepared
CC from the RNA of peripheral blood lymphocytes of an HIV-1 infected
CC donor. A claimed pharmaceutical preparation comprises a carrier
CC and an antibody that includes the RSVF2-5 VH3 CDR3 (AAW16586) and
CC which may also include the VH3 CDR2 (AAW16584) and/or CDR1 (AAW16582)
CC or the entire Fd region (AAW16580), or is an Fab fragment and includes
CC the RSVF2-5 VL6 CDR3 (AAW16594), CDR2, CDR1 (AAW16590) or entire
CC light chain (AAW16588). The preparation alternatively comprises a
CC carrier and a vector that includes a nucleotide sequence encoding
CC the antibody. The preparations can be used for the treatment or
CC prophylaxis of active RSV disease or infection (claimed), and may
CC also be used for RSV detection. The antibody binds and neutralises
CC antigenic subgroups A and B of RSV with high efficiency.

SQ Sequence 7 AA;

Query Match 89.5%; Score 34; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORP 6
| | | | |
| | | | |
Db 1 EDNORP 6

RESULT 8
ABBI2331
ID ABB12331 standard; Protein; 101 AA.
XX
AC ABB12331;
XX

DT 15-JAN-2002 (first entry)

DE Human bone marrow expressed protein SEQ ID NO: 85.

XX Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;
KM antinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
KM antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
KM antidiabetic; antidiabetic; antidiabetic; antiallergic;
KM immunostimulant; analgesic; cerebroprotective; antineoplastic; infection;
KM nervous system disorder; autoimmune disorder; inflammation; allergy.

OS Homo sapiens.

XX WO200174836-A1.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US10472.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PR 23-AUG-2000; 2000US-0649267.

PR 30-NOV-2000; 2000US-250583P.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Ford JR, Boyle BJ;

DR WPI: 2001-626375/72.

DR N-PSDB; ABA09575.

XX New bone marrow-expressed nucleic acids and polypeptides, useful for

PT diagnosis, treatment of inflammatory, autoimmune, neurological, cancer

PT and increasing hematopoiesis, stem cell survival and bone growth and

PT remodeling

XX Claim 10; Page 200-201; 380pp; English.

PS The present invention relates to bone marrow expressed polynucleotides

CC and proteins. These sequences can be used in the treatment of
CC inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
CC and peripheral nervous system diseases and neuropathies, such as
CC Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
CC disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
CC cell disorders, platelet disorders, stem cell disorders, bone
CC degenerative disorders, autoimmune disorders, for example multiple
CC sclerosis, diabetes and arthritis, viral and bacterial infections,
CC allergies and blood coagulation disorders. The present sequence is a
CC protein of the invention.

SQ Sequence 101 AA;

Query Match 89.5%; Score 34; DB 22; Length 101;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
| | | | |
| | | | |
Db 25 EDNKRPS 31

RESULT 9
AARS4054
ID AARS4054 standard; Protein; 106 AA.
XX
AC AARS4054;
XX

DT 25-MAR-2003 (updated)

DT 08-NOV-1994 (first entry)

DE Sequence of the VL region of monoclonal antibody MD3-4 against

DE hepatitis B virus surface antigen.

KM Hepatitis B virus; surface antigen; monoclonal antibody; therapy;

KM HBsAg; diagnosis; HBV.

XX Synthetic.

FH Key Location/Qualifiers

FT Region 1..95

FT Region /label= V III

FT Region /label= CDR 1

FT Region 50..56

FT Region /label= CDR 2

FT Region 88..96

FT Region /label= CDR 3

FT Region 95..106

FT Region /label= J 3

PN WO9411495-A1.

PD 26-MAY-1994.

PF 06-NOV-1992; 92WO-US09749.

PR 06-NOV-1992; 92WO-US09749.

XX (SANO) SANDOZ LTD.

PI Ostberg LG;

DR WPI, 1994-183497/22.

DR N-PSDB; AAQ64057.

XX Monoclonal antibodies active against Hepatitis B surface antigen

PT - for diagnosis and treatment of Hepatitis B virus

XX Example; Page 43; 53pp; English.

PS Human volunteers were immunised with hepatitis B vaccine. MD3-4, ZM1-

CC 2, ZM1-1 and PEl-1 hybridoma cell lines were derived from


```

PR 05-SEP-1986; 86US-0904517.
PR 31-OCT-1986; 86US-0925196.
PR 11-MAY-1988; 88US-0192754.
PR 15-JUN-1990; 90US-0538796.
PR 27-MAR-1991; 91US-0676036.
PR 14-JUN-1994; 94US-0259372.
PR 06-JUN-1995; 95US-0468671.
XX
XX (SANO ) SANDOZ LTD.
XX
XX Ostberg LG;
XX
XX WPI: 1997-372021/34.
XX
XX N-PSDB: AAT85845.
XX
XX
XX Treatment of hepatitis B - with human monoclonal antibody
XX
XX
XX Example 8; Column 23-24; 25pp; English.
XX
XX
XX This is the amino acid sequence of the light chain variable (VL) region
XX from the human monoclonal antibody (MAb) MD3-4. The MAb was generated
XX by immunising humans with a hepatitis B virus (HBV) vaccine, isolating
XX peripheral blood lymphocytes (PBL) and fusing them with a mouse/human
XX xenogeneic cell line SP2-4. 5 cell lines were isolated: pbl-1, ZM1-1,
XX ZM1-2, MD3-4 and L03-3. The cell lines were then tested for production
XX of an anti-hepatitis B virus surface antigen antibody by ELISA. The Mabs
XX are then purified from large scale cell culture by protein A
XX chromatography, size separation on Sephacryl S300 gel and ion exchange
XX chromatography on Q-Sepharose. The heavy and light chains of the Mabs
XX were isolated and their amino acid sequences determined. Primers were
XX generated and used to amplify cDNA synthesised from RNA purified from
XX each hybridoma cell line. The sequences of the heavy and light chains
XX (nucleic acid and amino acid) from Mabs PBL-1, ZM1-1, ZM1-2 and MD3-4
XX are shown in AAT8588-45 and AAM24984-91. The Mabs can be used to treat
XX HBV infections in immunosuppressed patients or patients with chronic
XX active hepatitis, especially liver transplant patients.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX
XX Sequence 106 AA;
XX
XX
XX Query Match 89.5%; Score 34; DB 18; Length 106;
XX Best Local Similarity 85.7%; Pred. No. 22;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EDNORPS 7
XX |||:|||
XX 49 EDNKRPS 55
XX
XX
XX RESULT 12
XX AAM16588
XX ID AAM16588 standard; Protein; 109 AA.
XX
XX AAM16588;
XX
XX 30-NOV-1997 (first entry)
XX
XX DE Anti-RSV F glycoprotein antibody RSVF2-5 light chain VL6.
XX
XX
XX Respiratory syncytial virus; RSV; monoclonal antibody; CDR;
XX complementarity determining region; pneumonia; bronchiolitis;
XX diagnosis; therapy; vaccine; RSVF2-5.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT 1..19
XX FT Peptide /label= FR1
XX /note= "framework region 1"
XX FT 20..32
XX FT Region /label= CDR1
XX /note= "complementarity determining region 1
XX (Claim 9)"
XX

```

FT	Region	33...47	/label= FR2	
FT		/note= "framework region 2"		
FT	Region	48...54	/label= CDR2	
FT		/note= "complementarity determining region 2 (Claim 8) "		
FT	Region	55...88	/label= FR3	
FT		/note= "framework region 3"		
FT	Region	89...98	/label= CDR3	
FT		/note= "complementarity determining region 3 (Claim 7) "		
FT	Region	99...109	/label= FR4	
FT		/note= "framework region 4"		
XX				
PN	WO9710846-A1.			
PD	27-MAR-1997.			
XX				
PF	18-SEP-1996;	96WO-US14937.		
XX				
PR	18-SEP-1995;	95US-0003931.		
XX				
PA	(INTR-) INTRACEL CORP.			
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.			
XX				
PI	Chanock RM, Crowe JE, Gilmour PS, Murphy BR, Pilkington GR;			
DR	WPI; 1997-202621/18.			
DR	N-PSDB; AAT66557.			
XX				
PT	Composition comprising respiratory syncytial virus antibody - useful for treatment or prophylaxis of active disease or infection			
XX				
PS	Claim 10; Page 52-53; 71pp; English.			
XX				
XX	This polypeptide sequence comprises the light chain VL6 region of a novel neutralising human monoclonal antibody to respiratory syncytial virus (RSV). The antibody, designated RSVF2-5 (ATCC 69909), selectively binds to an RSV F glycoprotein epitope. cDNA encoding the RSVF2-5 Fd VH3 and light chain VL6 (see AAT6556-57) was isolated from a phage library prepared from RNA extracted from the peripheral blood lymphocytes of an HIV-1 infected donor. A claimed pharmaceutical preparation comprises a carrier and an antibody that includes the RSVF2-5 VH3 CDR3 (AAW16586) and which may also include the VH3 CDR2 (AAW16584) and/or CDR1 or the entire Fd Fd region (AAW16580), or is an Fab fragment and further includes the RSVF2-5 VL6 CDR3 (AAW16594), CDR2 (AAW16592), CDR1 (AAW16590) or entire light chain (AAW16588). The preparation alternatively comprises a carrier and a vector that includes a nucleotide sequence encoding the antibody. The preparations can be used for the treatment or prophylaxis of active RSV disease or infection (claimed), and may also be used for RSV detection. The antibody binds and neutralises antigenic subgroups A and B of RSV with high efficiency.			
XX				
SO	Sequence	109 AA;		
	Query Match	89.5%; Score 34; DB 18; Length 109;		
	Best Local Similarity	100.0%; Pred. No. 23;		
	Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 EDNORP 6			
DB	48 EDNORP 53			
	RESULT 13			
	AAW14785			
ID	AAW14785 standard; Protein; 109 AA.			
XX				

AC	AAM14785;
XX	
DT	20-JUN-1997 (first entry)
DE	Human monoclonal antibody RSVF2-5 light chain VL6.
XX	
KW	Respiratory syncytial virus; RSV; monoclonal antibody; Mab;
XX	
KW	diagnosis; prophylaxis; immunotherapy; therapy; Fab; Fd.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..19
FT	/label= FR1
FT	20..32
FT	/label= CDR1
FT	/note= "Claim 9"
FT	33..47
FT	/label= FR2
FT	48..54
FT	/label= CDR2
FT	/note= "Claim 8"
FT	55..88
FT	/label= FR3
FT	89..98
FT	/label= CDR3
FT	/note= "Claim 7"
FT	99..109
FT	/label= FR4
XX	
PN	WO9711177-A1.
PD	
PD	27-MAR-1997.
XX	
PF	18-SEP-1996;
XX	
PR	96WO-US14944.
XX	
PR	18-SEP-1995;
XX	
PA	95US-0003931.
PI	(INTR-) INTRACEL CORP.
XX	
PI	Gilmour PS, Pilkington GR;
XX	
DR	WPI: 1997-202886/18.
XX	
DR	N-PsDB; AAT63418.
PT	
PT	Monoclonal antibody specific for respiratory syncytial virus - used
PT	for diagnosis and immuno-prophylaxis or immuno-therapy of RSV
PT	disease
XX	
PS	Claim 10; Page 49-50; 68pp; English.
XX	
CC	The light chain variable region (AAM14785) and heavy chain variable
CC	region (AAM14784) are provided of a novel, fully human monoclonal
CC	antibody (MAb), designated RSVF2-5, which selectively binds to
CC	the F glycoprotein of respiratory syncytial virus (RSV) and which
CC	neutralises RSV in vivo. To obtain RSVF2-5, a packaged phage
CC	library was prep'd. from amplified heavy and light chain variable
CC	region clones derived from the peripheral blood lymphocytes of an
CC	HIV-1 infected donor. The phage library was panned with RSV
CC	proteins and isolated clones were sequenced. Isolated nucleic
CC	acids (AAT63417-18) can be used to produce novel polypeptides, esp.
CC	Fd or Fab fragments useful in the diagnosis of RSV disease, in
CC	methods for detecting the presence of RSV in a sample, and in the
CC	immunoprophylaxis and immunotherapy of RSV disease.
XX	
SQ	Sequence 109 AA;

```

Query Match      89.5%; Score 34; DB 18; Length 109;
Best Local Similarity 100.0%; Pred. NO. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 EDNRP 6
|||||

```

DB	48	EDNORP	53
RESULT 14			
XX	AA	AB53654	
XX	AC	AB53654	standard; Protein; 114 AA.
XX	AC	AB53654;	
XX	DT	09-MAR-2001	(first entry)
XX	DE	Human colon cancer antigen protein sequence	SEQ ID NO:1194.
KW	KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;	
KW	KW	identification; cytostatic; cardioactive; neuroprotective; vulnery;	
KW	KW	immunomodulatory; muscular; gynaecological; gastrointestinal;	
KW	KW	neurotropic; antiinfective; antibacterial; gene therapy; wound;	
KW	KW	neural disorder; immune system disorder; muscular disorder;	
KW	KW	reproductive disorder; gastrointestinal disorder; renal disorder;	
KW	KW	infectious disease; cardiovascular disorder.	
XX	OS	Homo sapiens.	
XX	PN	WO20055351-A1.	
XX	PD	21-SEP-2000.	
XX	PE	08-MAR-2000; 2000WO-US05883.	
XX	PR	12-MAR-1999; 99US-0124270.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Rosen CA, Ruben SM;	
XX	DR	WPI; 2000-587534/55.	
XX	DR	N-PSDB; AAC98411.	
XX	PT	Colon cancer associated gene sequences, referred to as colon cancer	
XX	PT	antigens, useful for the treatment, prevention, and diagnosis of colon	
XX	PT	disorders such as colon cancer -	
XX	PS	Claim 11; Page 1776; 2104pp; English.	
XX	XX	AAC97991 to AAC98763 encode the human colon cancer associated proteins,	
XX	XX	called human colon cancer antigens, given in AB53234 to AB54006. The	
XX	XX	human colon cancer antigens can have cyrostatic, cardioactive, muscular;	
XX	XX	neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and	
XX	XX	vulnery, neurotropic, antiinfective and antibacterial activities, and	
XX	XX	can be used in gene therapy. The colon cancer antigen polynucleotides,	
XX	XX	proteins and antibodies to the proteins are useful for the prevention,	
XX	XX	treatment and diagnosis of colon disorders, such as colon cancer. The	
XX	XX	polynucleotides may be used in diagnostics and research, such as for	
XX	XX	chromosome identification, and as hybridisation probes. The proteins	
XX	XX	may also be used to prevent diseases such as neural disorders, immune	
XX	XX	system disorders, muscular disorders, reproductive disorders,	
XX	XX	gastrointestinal disorders, wounds, renal disorders, infectious	
XX	XX	diseases, and cardiovascular disorders. AAC98764 to AAC98772 and	
XX	XX	AB54007 represent sequences used in the exemplification of the present	
XX	XX	invention.	
XX	XX	Sequence 114 AA;	
XX	XX		

```

Query Match          89.5%;   Score 34;  DB 21;  Length 114;
Best Local Similarity 85.7%;   Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDNQRP 7
      |||:||||
Db      37 EDNKRPS 43

RESULT 15

```

AAU82012
 ID AAU82012 standard; Protein; 233 AA.
 XX
 AC AAU82012;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human secreted protein SECP38.
 XX
 KW Human; SECP; antiinflammatory; cytosolic; cardiant;
 KW immunosuppressive; antiviral; anti-HIV; antiarthritic; antirheumatic;
 KW muscular active general; anticonvulsant; nootropic; neuroprotective;
 KW antiallergic; hypotensive; cardiovascular disorder; atherosclerosis;
 KW hypertension; myocardial infarction; autoimmune disorder;
 KW inflammatory disorder; AIDS; acquired immunodeficiency syndrome;
 KW allergy; rheumatoid arthritis; cell proliferative disorder; cancer;
 KW developmental disorder; Duchenne muscular dystrophy;
 KW neurological disorder; epilepsy; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 PN MO200198353-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 20-JUN-2001; 2001WO-US19862.
 XX
 PR 20-JUN-2000; 2000US-212890P.
 PR 23-JUN-2000; 2000US-213466P.
 PR 27-JUN-2000; 2000US-214601P.
 PR 31-JUL-2000; 2000US-222372P.
 PR 08-SEP-2000; 2000US-221435P.
 PR 15-SEP-2000; 2000US-232889P.
 XX
 PA (INCY-) INCTE GENOMICS INC.
 XX
 PI Hillman JL, Tang YT, Yue H, Elliott VS, Tribouley CM, Lee EA;
 PI Ramkumar J, Lal P, Xu Y, Warren BA, Hafalia AJA, Baughn MR;
 PI Azimzai Y, Bateria S, Burford N, Yao MG, Nguyen DB, Lu DAM;
 PI Walla NK, Gandhi AR, Au-Young J, Patterson C;
 XX
 DR WPI; 2002-090431/12.
 DR N-PSDB; ABK28671.
 XX
 PT Forty four human secreted proteins (referred to as SECP-1 to SECP-44),
 PT useful in the diagnosis, treatment and prevention of cardiovascular
 PT (e.g. atherosclerosis), autoimmune/inflammatory (e.g. allergies) and
 PT cell proliferative disorders -
 XX
 PS Claim 1; Page 158-159; 195pp; English.
 XX
 CC The invention relates to forty four human secreted proteins (referred to
 CC as SECP-1 to SECP-44) and the nucleic acids encoding them. Also
 CC included are a host cell transformed with the nucleic acid, a
 CC transgenic animal comprising the nucleic acid, an anti-SECP
 CC antibody, use of the SECP proteins in isolating agonists and antagonists
 CC of SECP activity and a method of isolating compounds which alter the
 CC expression of the SECP nucleic acid. The SECP polynucleotides and
 CC polypeptides are useful in the diagnosis, treatment and prevention of
 CC cardiovascular (e.g. atherosclerosis, hypertension, myocardial
 CC infarction), autoimmune/inflammatory (e.g. acquired immunodeficiency
 CC syndrome (AIDS), allergies, rheumatoid arthritis), cell proliferative
 CC (e.g. cancer), developmental (e.g. Duchenne and Becker muscular
 CC dystrophy), and neurological (e.g. epilepsy, Alzheimer's disease)
 CC disorders. Numerous other examples of each disorder are given in the
 CC specification. The present sequence represents a SECP protein.
 CC
 SO Sequence 233 AA;

Qy 1 EDNRPS 7
 |||:||||
 Db 68 EDNRPS 74

Search completed: January 12, 2004, 06:58:57
 Job time : 17.9531 secs

Query Match 89.5%; Score 34; DB 23; Length 233;
 Best Local Similarity 85.7%; Pred. No. 51;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 07:02:25 ; Search time 14 Seconds

(without alignments)
100.755 Million cell updates/sec

Title: US-09-829-495-65

Perfect score: 38

Sequence: 1 EDNQRPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	9	US-09-832-312-65
2	38	100.0	7	11	US-09-972-656-13
3	38	100.0	7	12	US-09-829-495-65
4	38	100.0	98	12	US-10-308-817-117
5	38	100.0	103	11	US-09-972-656-110
6	38	100.0	110	15	US-10-269-805-42
7	38	100.0	110	15	US-10-269-805-52
8	38	100.0	110	15	US-10-269-805-58
9	38	100.0	161	15	US-10-106-698-6785
10	38	100.0	216	11	US-09-972-656-108
11	38	100.0	258	11	US-09-880-748-1234
12	38	100.0	114	9	US-09-925-299-1194
13	38	100.0	114	11	US-09-925-299-1194
14	38	100.0	243	12	US-10-322-673-55
15	38	100.0	244	11	US-09-880-748-1842

16	33	86.8	7	11	US-09-972-656-14	Sequence 14, Appl
17	33	86.8	7	11	US-09-972-656-20	Sequence 20, Appl
18	33	86.8	216	11	US-09-972-656-90	Sequence 90, Appl
19	33	86.8	218	11	US-09-972-656-102	Sequence 102, Appl
20	33	86.8	248	11	US-09-880-748-1273	Sequence 1273, Appl
21	33	86.8	252	11	US-09-880-748-1128	Sequence 1128, Appl
22	33	86.8	253	11	US-09-880-748-1227	Sequence 1227, Appl
23	33	86.8	259	11	US-09-880-748-1277	Sequence 1277, Appl
24	32	84.2	110	15	US-10-269-805-50	Sequence 40, Appl
25	32	84.2	110	15	US-10-269-805-50	Sequence 50, Appl
26	32	84.2	245	11	US-09-880-748-1714	Sequence 1714, Appl
27	31	81.6	7	11	US-09-563-222-24	Sequence 24, Appl
28	31	81.6	235	12	US-10-120-414-82	Sequence 82, Appl
29	31	81.6	235	12	US-10-120-414-83	Sequence 83, Appl
30	31	81.6	248	11	US-09-880-748-1965	Sequence 1965, Appl
31	31	81.6	248	11	US-09-880-748-1974	Sequence 1974, Appl
32	31	81.6	249	12	US-10-139-785-53	Sequence 53, Appl
33	31	81.6	249	13	US-10-039-785-53	Sequence 53, Appl
34	31	81.6	250	11	US-09-880-748-1540	Sequence 1540, Appl
35	31	81.6	253	11	US-09-880-748-1488	Sequence 1488, Appl
36	31	81.6	384	15	US-10-156-761-14623	Sequence 14623, A
37	30	78.9	7	11	US-09-791-153A-8	Sequence 8, Appl
38	30	78.9	108	11	US-09-791-153A-74	Sequence 74, Appl
39	30	78.9	109	14	US-10-025-687-12	Sequence 12, Appl
40	30	78.9	109	15	US-10-001-934-38	Sequence 38, Appl
41	30	78.9	109	15	US-10-001-934-42	Sequence 42, Appl
42	30	78.9	109	15	US-10-001-934-44	Sequence 44, Appl
43	30	78.9	109	15	US-10-001-934-46	Sequence 46, Appl
44	30	78.9	109	15	US-10-001-934-48	Sequence 48, Appl
45	30	78.9	109	15	US-10-001-934-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-09-832-312-65
; Sequence 65, Application US/09832312
; Patent No. US20010049829A1
; GENERAL INFORMATION:
; APPLICANT: Busfield et al
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 65
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-65
Query Match 100.0%; Score 38; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 EDNQRPS 7
Db 1 EDNQRPS 7
RESULT 2
US-09-972-656-13
; Sequence 13, Application US/09972656

```
Publication No. US20030099647A1
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
APPLICANT: Tsai, Mei-Mei
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 13
LENGTH: 7
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-13
```

```
Query Match      100.0%; Score 38; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 EDNORPS 7
        |||||
Db       1 EDNORPS 7
```

RESULT 3

```
US-09-829-495-65
Sequence 65, Application US/09829495
Publication No. US20040001826A1
GENERAL INFORMATION:
APPLICANT: Busfield SJ
APPLICANT: Valleval J
APPLICANT: Jandrot-Petruis M
APPLICANT: Vainchenker W
APPLICANT: Gili DS
APPLICANT: Qian MD
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-234
CURRENT APPLICATION NUMBER: US/09/829,495
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 09/610,118
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/503,387
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 09/454,824
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 09/345,468
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 65
LENGTH: 7
TYPE: PRT
ORGANISM: Homo sapiens
US-09-829-495-65
```

```
Query Match      100.0%; Score 38; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 EDNORPS 7
        |||||
Db       1 EDNORPS 7
```

RESULT 4

```
US-10-308-817-117
Sequence 117, Application US/10308817
Publication No. US20030219861A1
GENERAL INFORMATION:
APPLICANT: Rother, Russell
APPLICANT: Wu, Dayang
```

```
TITLE OF INVENTION: HYBRID ANTIBODIES
FILE REFERENCE: 1087-37
CURRENT APPLICATION NUMBER: US/10/308,817
CURRENT FILING DATE: 2002-12-03
NUMBER OF SEQ ID NOS: 195
SOFTWARE: PatentIn version 3.1
SEQ ID NO 117
LENGTH: 98
TYPE: PRT
ORGANISM: human
US-10-308-817-117
```

```
Query Match      100.0%; Score 38; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 EDNORPS 7
        |||||
Db       51 EDNORPS 57
```

RESULT 5

```
US-09-972-656-130
Sequence 130, Application US/09972656
Publication No. US20030099647A1
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
APPLICANT: Tsai, Mei-Mei
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 130
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Misc.
LOCATION: (7)..(7)
OTHER INFORMATION: Unidentified
NAME/KEY: Misc.
LOCATION: (33)..(33)
OTHER INFORMATION: Unidentified
NAME/KEY: Misc.
LOCATION: (34)..(34)
OTHER INFORMATION: Unidentified
NAME/KEY: Misc.
LOCATION: (35)..(35)
OTHER INFORMATION: Unidentified
NAME/KEY: Misc.
LOCATION: (36)..(36)
OTHER INFORMATION: Unidentified
US-09-972-656-130
```

```
Query Match      100.0%; Score 38; DB 11; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 EDNORPS 7
        |||||
Db       56 EDNORPS 62
```

RESULT 6

```
US-10-269-805-42
Sequence 42, Application US/10269805
Publication No. US20030124129A1
GENERAL INFORMATION:
APPLICANT: OLINER, JONATHAN D.
TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
```

FILE REFERENCE: A-722
CURRENT APPLICATION NUMBER: US/10/269,805
CURRENT FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: US 60/328,604
PRIOR FILING DATE: 2001-10-11
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-10-269-805-42

Query Match 100.0%; Score 38; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
|||||
DB 51 EDNORPS 57

RESULT 7
US-10-269-805-52
Sequence 52, Application US/10269805
Publication No. US20030124129A1
GENERAL INFORMATION:
APPLICANT: OLINER, JONATHAN D.
TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
FILE REFERENCE: A-722
CURRENT APPLICATION NUMBER: US/10/269,805
CURRENT FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: US 60/328,604
PRIOR FILING DATE: 2001-10-11
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-10-269-805-52

Query Match 100.0%; Score 38; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
|||||
DB 51 EDNORPS 57

RESULT 8
US-10-269-805-58
Sequence 58, Application US/10269805
Publication No. US20030124129A1
GENERAL INFORMATION:
APPLICANT: OLINER, JONATHAN D.
TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
FILE REFERENCE: A-722
CURRENT APPLICATION NUMBER: US/10/269,805
CURRENT FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: US 60/328,604
PRIOR FILING DATE: 2001-10-11
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn version 3.1
SEQ ID NO 58
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-10-269-805-58

Query Match 100.0%; Score 38; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORPS 7
|||||
DB 51 EDNORPS 57

RESULT 9
US-10-106-698-6785
Sequence 6785, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA00521
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 6785
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (149)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6785

Query Match 100.0%; Score 38; DB 15; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
|||||
DB 72 EDNORPS 78

RESULT 10
US-09-972-656-108
Sequence 108, Application US/09972656
Publication No. US2003009647A1
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
APPLICANT: Tsai, Mei-Mei
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 108
LENGTH: 216
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-108

Query Match 100.0%; Score 38; DB 11; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
|||||
DB 51 EDNORPS 57

RESULT 11

US-09-880-748-1234
; Sequence 1234, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Bly's
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1234
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1234

Query Match 100.0%; Score 38; DB 11; Length 258;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
|||:||||
DB 198 EDNRPS 204

RESULT 12

US-09-925-299-1194
; Sequence 1194, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1194
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1194

Query Match 89.5%; Score 34; DB 9; Length 114;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
|||:||||
DB 37 EDNRPS 43

RESULT 13

US-09-925-299-1194
; Sequence 1194, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1194
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1194

Query Match 89.5%; Score 34; DB 11; Length 114;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
|||:||||
DB 37 EDNRPS 43

RESULT 14

US-10-322-673-55
; Sequence 55, Application US/10322673
; Publication No. US20030180296A1
; GENERAL INFORMATION:
; APPLICANT: Salcedo et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to TRAIL
; FILE REFERENCE: PF585
; CURRENT APPLICATION NUMBER: US/10/322,673
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: 60/341,237
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/369,877
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,828
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/396,591
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/403,370
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/425,737
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 55
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: CM075A01 scFv
US-10-322-673-55

Query Match 89.5%; Score 34; DB 12; Length 243;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
|||:||||
DB 183 EDNRPS 189

RESULT 15

US-09-880-748-1842

; Sequence 1842; Application US/09880748
; Publication No. US20030059937A1

GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523

; CURRENT APPLICATION NUMBER: US/09/880,748

; PRIOR APPLICATION NUMBER: 2001-06-15

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 3239

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 1842

; LENGTH: 244

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-880-748-1842

Query Match

Best Local Similarity 89.5%; Score 34; DB 11; Length 244;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7

Db 186 EDNRPS 192

Search completed: January 12, 2004, 07:19:56
Job time : 14 secs

THIS PAGE BLANK (USPTO)

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-649-991-91

Query Match 56.5%; Score 35; DB 2; Length 132;
Best Local Similarity 54.5%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNY 11
|||::||
Db 122 TRSSSQVSONY 132

RESULT 14
US-08-649-991-92
Sequence 92, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
TITLE OF INVENTION: MATERNOCETAL TRANSMISSION OF HIV-1
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7176
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 92:

SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-649-991-92

Query Match 56.5%; Score 35; DB 2; Length 132;
Best Local Similarity 54.5%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNY 11
|||::||
Db 122 TRSSSQVSONY 132

RESULT 15
US-09-071-035-344
Sequence 344, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 344:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-344

Query Match 56.5%; Score 35; DB 4; Length 431;
Best Local Similarity 38.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
||:|::|:
Db 268 TKENSKIANNYIE 280

Search completed: January 12, 2004, 07:04:01
Job time : 12.7656 secs

QY 1 TRSSGSIASNYV 12
|||
Db 101 TRKEGYIPSNYV 112

RESULT 10

US-09-446-301A-49
; Sequence 49, Application US/09446301A
; Patent No. 6506893
; GENERAL INFORMATION:
; APPLICANT: EL SOLH, NEVINE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; FILE REFERENCE: 03715-0059
; CURRENT APPLICATION NUMBER: US/09/446,301A
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 49
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Scaphylococcus sp.
US-09-446-301A-49

Query Match 58.1%; Score 36; DB 4; Length 522;
Best Local Similarity 38.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 13
|||
Db 79 TRSGGEVTNRYR 91

RESULT 11
US-09-099-932-49
; Sequence 49, Application US/09099932
; Patent No. 6570001
; GENERAL INFORMATION:
; APPLICANT: EL SOLH, NEVINE
; APPLICANT: ALLIGRET, JEANINE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; FILE REFERENCE: 03495.0173-00000
; CURRENT APPLICATION NUMBER: US/09/099,932
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/050,380
; EARLIER FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 49
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Staphylococcus
US-09-099-932-49

Query Match 58.1%; Score 36; DB 4; Length 522;
Best Local Similarity 38.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 13
|||
Db 79 TRSGGEVTNRYR 91

RESULT 12
US-08-630-915A-131
; Sequence 131, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOMLICKS, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 131:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-915A-131

Query Match 56.5%; Score 35; DB 4; Length 58;
Best Local Similarity 58.3%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
|||
Db 44 TRKEGFIPSNYV 55

RESULT 13
US-08-649-991-91
; Sequence 91, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narva, Remy
; APPLICANT: Rogues, Pierre
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOPETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
ADDRESSER: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,509A
FILING DATE: 21-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/232,545
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-0074-999
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-426-509A-17

Query Match 61.3% Score 38; DB 4; Length 505;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
DB 101 TRREGYIPSNYV 112

RESULT 8
US-08-232-545-17
Sequence 17, Application US/08232545
Patent No. 6506578
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Gishizky, Michael
APPLICANT: Sures, Irmah G.
TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,545
FILING DATE: 22-Apr-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-050
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-232-545-17

Query Match 61.3% Score 38; DB 4; Length 505;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
DB 101 TRREGYIPSNYV 112

RESULT 9
PCT-US95-05008-17
Sequence 17, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-Apr-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-Apr-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-05008-17

Query Match 61.3% Score 38; DB 5; Length 505;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-988-925-16

Query Match 62.9%; Score 39; DB 1; Length 110;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
| | | | | : | | | | |
Db 23 TLSSGNIENNYV 34

RESULT 5
US-08-362-780-16
Sequence 16, Application US/08362780
Patent No. 5968509

GENERAL INFORMATION:
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: Antibody Preparation
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye pc
STREET: 8th Floor, 1100 No. 5968509th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,780
FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/862,543
FILING DATE: 23-JUNE-1992
APPLICATION NUMBER: GB 9021679.7
FILING DATE: 05-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB91/01726
FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-780-16

Query Match 62.9%; Score 39; DB 2; Length 110;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
| | | | | : | | | | |
Db 23 TLSSGNIENNYV 34

RESULT 6
US-08-362-780-26

Sequence 26, Application US/08362780
Patent No. 5968509

GENERAL INFORMATION:
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: Antibody Preparation
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye pc
STREET: 8th Floor, 1100 No. 5968509th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,780
FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/862,543
FILING DATE: 23-JUNE-1992
APPLICATION NUMBER: GB 9021679.7
FILING DATE: 05-OCT-1990
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/GB91/01726
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-780-26

Query Match 62.9%; Score 39; DB 2; Length 110;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
| | | | | : | | | | |
Db 23 TLSSGNIENNYV 34

RESULT 7
US-08-426-509A-17
Sequence 17, Application US/08426509A
Patent No. 6326469

GENERAL INFORMATION:
APPLICANT: Ullrich,, Axel
APPLICANT: Gishizky,, Mikhail
APPLICANT: Sures,, Irman G.
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA

CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988,925
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206422.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB92/01933
FILING DATE: 21-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-988-925-4

Query Match 62.9%; Score 39; DB 1; Length 13;
Best Local Similarity 66.7%; Pred. No. 0.42;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
DB 1 TLSSGNIENNYV 12

RESULT 3
US-08-362-780-4
Sequence 4, Application US/08362780
Patent No. 5968509
GENERAL INFORMATION:
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: Antibody Preparation
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye pc
STREET: 8th Floor, 1100 No. 5968509th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,780
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/862,543
FILING DATE: 23-JUNE-1992
APPLICATION NUMBER: GB 9021679.7

FILING DATE: 05-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB91/01726
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-780-4

Query Match 62.9%; Score 39; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 0.42;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
DB 1 TLSSGNIENNYV 12

RESULT 4
US-07-988-925-16
Sequence 16, Application US/07988925
Patent No. 5585097
GENERAL INFORMATION:
APPLICANT: Bolt, Sarah L
APPLICANT: Clark, Michael R
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: Antibody preparation
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye pc
STREET: 11th Floor, 1100 No. 5585097th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988,925
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206422.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB92/01933
FILING DATE: 21-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:56:19 ; Search time 10.7656 Seconds
(without alignments)
51.092 Million cell updates/sec

Title: US-09-829-495-64
Perfect score: 62
Sequence: 1 TRSSGSIASNYQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100.0	13	4	US-09-832-312-64	Sequence 64, Appl
2	39	62.9	13	1	US-07-988-925-4	Sequence 4, Appl
3	39	62.9	13	1	US-08-362-780-4	Sequence 4, Appl
4	39	62.9	110	1	US-07-988-925-16	Sequence 16, Appl
5	39	62.9	110	2	US-08-362-780-16	Sequence 16, Appl
6	39	62.9	110	2	US-08-362-780-26	Sequence 26, Appl
7	38	61.3	505	4	US-08-426-509A-17	Sequence 17, Appl
8	38	61.3	505	4	US-08-426-509A-17	Sequence 17, Appl
9	38	61.3	505	5	PCT-US95-05008-17	Sequence 17, Appl
10	36	58.1	522	4	US-09-446-301A-49	Sequence 49, Appl
11	36	58.1	522	4	US-09-099-932-49	Sequence 49, Appl
12	35	56.5	58	4	US-08-630-915A-131	Sequence 131, Appl
13	35	56.5	132	2	US-08-649-991-91	Sequence 91, Appl
14	35	56.5	132	2	US-08-649-991-92	Sequence 92, Appl
15	35	56.5	431	4	US-09-071-035-344	Sequence 344, Appl
16	35	56.5	474	4	US-09-071-035-342	Sequence 342, Appl
17	35	56.5	512	4	US-08-426-509A-16	Sequence 16, Appl
18	35	56.5	512	4	US-08-232-545-16	Sequence 16, Appl
19	35	56.5	512	4	PCT-US95-05008-16	Sequence 16, Appl
20	35	56.5	1129	4	US-09-252-991A-22330	Sequence 22330, A
21	34	54.8	29	3	US-09-070-504-1	Sequence 1, Appl
22	34	54.8	30	3	US-09-070-504-2	Sequence 2, Appl
23	34	54.8	37	1	US-08-233-389C-9	Sequence 9, Appl
24	34	54.8	37	2	US-08-801-863-9	Sequence 9, Appl
25	34	54.8	37	2	US-08-486-596A-9	Sequence 9, Appl
26	34	54.8	37	2	US-09-004-713-9	Sequence 9, Appl
27	34	54.8	37	2	US-08-619-841-2	Sequence 2, Appl

28	34	54.8	37	2	US-08-446-929A-2	Sequence 2, Appl
29	34	54.8	37	3	US-09-070-504-3	Sequence 3, Appl
30	34	54.8	37	3	US-09-070-504-8	Sequence 8, Appl
31	34	54.8	37	3	US-09-070-504-9	Sequence 9, Appl
32	34	54.8	37	6	5424221-4	Patent No. 5424221
33	34	54.8	37	1	US-07-776-272-20	Sequence 20, Appl
34	34	54.8	78	1	US-09-461-325-222	Sequence 222, Appl
35	34	54.8	98	1	US-08-211-202-111	Sequence 111, Appl
36	34	54.8	98	2	US-08-665-202-38	Sequence 38, Appl
37	34	54.8	98	4	US-09-315-574-38	Sequence 38, Appl
38	34	54.8	107	4	US-08-635-109-5	Sequence 5, Appl
39	34	54.8	109	1	US-08-276-852-147	Sequence 147, Appl
40	34	54.8	109	1	US-08-899-575-147	Sequence 147, Appl
41	34	54.8	109	1	US-08-899-575-147	Sequence 147, Appl
42	34	54.8	109	3	US-08-240-274-59	Sequence 59, Appl
43	34	54.8	109	4	US-09-025-769B-32	Sequence 32, Appl
44	34	54.8	109	4	US-09-025-769B-51	Sequence 51, Appl
45	34	54.8	109	5	PCT-US95-08743-147	Sequence 147, Appl

ALIGNMENTS

```
RESULT 1
US-09-832-312-64
; Sequence 64, Application US/09832312
; Patent No. 6548741
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832.312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610.118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503.387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454.824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345.468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-64

Query Match      100.0%; Score 62; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TRSSGSIASNYQ 13
        |||||
Db       1 TRSSGSIASNYQ 13

RESULT 2
US-07-988-925-4
; Sequence 4, Application US/07988925
; Patent No. 5585097
; GENERAL INFORMATION:
; APPLICANT: Bolt, Sarah L
; APPLICANT: Clark, Michael R
; APPLICANT: Gorman, Scott D
; APPLICANT: Rutledge, Edward G
; APPLICANT: Waldmann, Herman
; TITLE OF INVENTION: antibody preparation
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Nixon and Vanderhye pc
; STREET: 11th Floor, 1100 No. 5585097th Glebe Road
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2004, 06:54:49 ; Search time 11.7912 seconds
(without alignments)
106.117 Million cell updates/sec

Title: US-09-829-495-64
Perfect score: 62
Sequence: 1 TRSSGSIASNYVQ 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	83.9	112	1 LGHU48	Ig lambda chain V-
2	52	83.9	131	1 LGHUB	Ig lambda chain pr
3	51	82.3	111	1 LGHUL	Ig lambda chain V-
4	43	69.4	111	1 LGHUST	Ig lambda chain V-
5	40	64.5	778	1 T39047	hypothetical prote
6	38	61.3	112	1 LGHUR	Ig lambda chain V-
7	38	61.3	505	1 TVHHC	protein-tyrosine k
8	38	61.3	556	2 B82379	conserved hypothet
9	38	61.3	606	2 T27072	hypothetical prote
10	38	61.3	726	2 A90771	hypothetical prote
11	38	61.3	726	2 E85633	hypothetical prote
12	38	61.3	726	2 C64839	probable ATPase YC
13	37	59.7	108	2 P50073	Ig kappa chain V r
14	37	59.7	236	2 E85098	hypothetical prote
15	37	59.7	1181	2 T30578	myosin IC - slime
16	36	58.1	140	2 H86292	F7H2.18 protein -
17	36	58.1	308	2 G90504	2-keto-3-deoxy glu
18	36	58.1	352	2 T27677	hypothetical prote
19	36	58.1	522	2 JCI204	vga protein - Step
20	36	58.1	539	2 B49114	protein-tyrosine k
21	36	58.1	549	2 S61706	RNA 5'-triphosphat
22	36	58.1	997	2 T15243	hypothetical prote
23	35	56.5	131	2 S24321	Ig lambda chain pr
24	35	56.5	278	2 T39517	dual-specificity M
25	35	56.5	402	2 C69110	glutamate N-acetyl
26	35	56.5	503	1 J01321	protein-tyrosine k
27	35	56.5	503	1 TVMSHC	protein-tyrosine k
28	35	56.5	512	1 TVHULY	protein-tyrosine k
29	35	56.5	537	2 I51592	protein-tyrosine k

30	35	56.5	541	2	S31645	protein-tyrosine k
31	35	56.5	574	2	T40213	hypothetical prote
32	35	56.5	788	2	AC2767	formate dehydrogen
33	35	56.5	788	2	P97547	probable oxidoredu
34	35	56.5	859	2	AE2217	hypothetical prote
35	35	56.5	1007	2	H81670	conserved hypothet
36	35	56.5	1167	2	T06146	disease resistance
37	34	54.8	47	2	B30607	Ig kappa chain V-I
38	34	54.8	55	2	P00829	E2/NS1 protein (cd
39	34	54.8	72	2	I37232	calcitonin gene-re
40	34	54.8	97	2	B89792	conserved hypothet
41	34	54.8	98	2	S36048	Ig lambda chain -
42	34	54.8	124	2	C48234	hypothetical prote
43	34	54.8	127	2	A25864	calcitonin gene-re
44	34	54.8	129	2	S78058	Ig lambda chain pr
45	34	54.8	130	2	S78057	Ig lambda chain pr

ALIGNMENTS

RESULT 1
LGHU48
Ig lambda chain V-VI region (Nig-48) - human
C:Species: Homo sapiens (man)
C:Date: 28-Feb-1980 #sequence_revision 28-Feb-1980 #text_change 02-Sep-1997
C:Accession: A01991
R.Takahashi, N.; Takayasu, T.; Isobe, T.; Shinoda, T.; Okuyama, T.; Shimizu, A.
J. Biochem. 86, 1523-1535, 1979
A>Title: Comparative study on the structure of the light chains of human immunoglobulins
A:Reference number: A01991; MUID:80094390; PMID:118171
A:Accession: A01991
A:Molecule type: protein
A:Residues: 1-112 <TRK>
A>Note: this is the first sequenced V region of lambda chain subgroup VI
C:Comment: This is a Bence Jones protein.
C:Genetics:
A:Gene: GDB:IGLV@
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer
F:15-93/Domain: immunoglobulin homology <IMM>
F:22-91/Disulfide bonds: #status predicted

Query Match 83.9%; Score 52; DB 1; Length 112;
Best Local Similarity 84.6%; Pred. No. 0.016;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
DB 23 TRTSDSIASNYVQ 35

RESULT 2
LGHUB
Ig lambda chain precursor V-VI region (EB4) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
C:Accession: A01990
R.Anderson, M.L.M.; Brown, L.; McKenzie, E.; Kellow, J.E.; Young, B.D.
Nucleic Acids Res. 13, 2931-2941, 1985
A>Title: Cloning and sequence analysis of an Ig lambda light chain mRNA expressed in the
A:Reference number: A01990; MUID:85215660; PMID:3923440
A:Accession: A01990
A:Molecule type: mRNA
A:Residues: 1-131 <AND>
C:Genetics:
A:Gene: GDB:IGLV@
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2

C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotrimer; immunoglobulin
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-131/Product: Ig lambda chain V-VI region (EB4) #status predicted <MAY>
 F:20-41/Region: framework 1
 F:14-112/Domain: immunoglobulin homology <IMM>
 F:42-54/Region: complementarity-determining 1
 F:55-69/Region: framework 2
 F:70-76/Region: complementarity-determining 2
 F:77-110/Region: framework 3
 F:111-118/Region: complementarity-determining 3
 F:119-131/Region: framework 4
 F:41-110/Disulfide bonds: #status predicted

Query Match 83.3%; Score 52; DB 1; Length 131;
 Best Local Similarity 84.6%; Pred. No. 0.019;
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 42 TNGSGSIASNYVQ 54

RESULT 3

LGHUR
 Ig lambda chain V-VI region (WLT) - human

C/Species: Homo sapiens (man)

C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997

C/Accession: A01989

R/Dwulet, F.E.; Strako, K.; Benson, M.D.

Scand. J. Immunol. 22, 653-660, 1985.

A/Title: Amino acid sequence of a lambda VI primary (AL) amyloid protein (WLT).

A/Reference number: A01989; PMID:86122667; PMID:4089539

A/Accession: A01989

A/Molecule type: Protein

A/Residues: 1-111 <DMU>

C/Genetics:

A/Gene: GDB:IGLV@

A/Cross-references: GDB:119342; OMIM:147240

A/Map position: 22q11.2-22q11.2

C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotrimer; immunoglobulin

F:1-22/Region: framework 1

F:15-93/Domain: immunoglobulin homology <IMM>

F:23-35/Region: complementarity-determining 1

F:36-50/Region: framework 2

F:51-57/Region: complementarity-determining 2

F:58-91/Region: framework 3

F:92-101/Region: complementarity-determining 3

F:102-111/Region: framework 4

F:22-91/Disulfide bonds: #status predicted

Query Match 82.3%; Score 51; DB 1; Length 111;
 Best Local Similarity 84.6%; Pred. No. 0.025;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 23 TNGSGSIASNYVQ 54

RESULT 4

LGHUR

Ig lambda chain V-VI region (SUT) - human

C/Species: Homo sapiens (man)

C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997

C/Accession: A01988

R/Solomon, A.; Kyle, R.A.; Frangione, B.

in Amyloidosis, Glenner, G.G., Osseman, E.F., Benditt, E.P., Calkins, E., Cohn, A.S., &

A/Title: Light chain variable region subgroups of monoclonal immunoglobulins in amyloidosis
 A/Reference number: A01988
 A/Accession: A01988
 A/Molecule type: protein
 A/Residues: 1-111 <SOL>

C/Genetics:

A/Gene: GDB:IGLV@

A/Cross-references: GDB:119342; OMIM:147240

A/Map position: 22q11.2-22q11.2

C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotrimer; immunoglobulin

F:1-22/Region: framework 1

F:15-93/Domain: immunoglobulin homology <IMM>

F:23-35/Region: complementarity-determining 1

F:36-50/Region: framework 2

F:51-57/Region: complementarity-determining 2

F:58-91/Region: framework 3

F:92-100/Region: complementarity-determining 3

F:101-111/Region: framework 4

F:22-91/Disulfide bonds: #status predicted

Query Match 69.4%; Score 43; DB 1; Length 111;
 Best Local Similarity 69.2%; Pred. No. 0.81;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 23 TRSDGSIAGYVQ 35

RESULT 5

TR9047

hypothetical protein SPAC6F6.13c - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 15-Sep-2000

C/Accession: T39047

R/Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A/Reference number: Z21797

A/Accession: T39047

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-778 <GEN>

A/Cross-references: EMBL:Z98981; PDB:1CAB11736.1; GSPDB:GN00066; SPDB:SPAC6F6.13c

A/Experimental source: strain 97zh-; coemid ccf6

C/Genetics:

A/Gene: SPDB:SPAC6F6.13c

A/Map position: 1

C/Superfamily: Candida albicans conserved hypothetical protein PLO9050.2

Query Match 64.5%; Score 40; DB 2; Length 778;
 Best Local Similarity 72.7%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SSGSIASNYVQ 13
 DB 39 SKGSITENYVQ 49

RESULT 6

LGHUR

Ig lambda chain V-VI region (AR) - human (tentative sequence)

C/Species: Homo sapiens (man)

C/Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 31-Mar-2000

C/Accession: A01987

R/Sletten, K.; Natvig, J.B.; Husby, G.; Juul, J.

Biochem. J. 195, 561-572, 1981

A/Title: The complete amino acid sequence of a prototype immunoglobulin-lambda light-chain

A/Reference number: A01987; PMID:82091000; PMID:6797401

A/Contents: amyloid protein AR

A/Accession: A01987

A:Molecule type: protein
 A:Residues: 1-112 <SLF>
 A:Note: about half of the lambda chain C region is missing from this protein
 C:Comment: This protein was isolated from the spleen of a patient with amyloidosis.
 C:Genetics:
 A:Gene: GDB:IGLV@
 A:Cross-references: GDB:119342; OMIM:147240
 A:Map position: 22q11.2-22q11.2
 C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into a C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: amyloid; heterotrimer; immunoglobulin
 F:15-93/Domain: immunoglobulin homology <IMM>
 F:22-91/Disulfide bonds: #status predicted

Query Match 61.3%; Score 38; DB 1; Length 112;
 Best Local Similarity 61.5%; Pred. No. 7.2;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 23 TSGSGSIADSFVQ 35

RESULT 7

TYVHHC
 protein-tyrosine kinase (EC 2.7.1.112) hck - human

C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1989 #sequence_revision 10-Nov-1995 #text_change 11-Jun-1999

C:Accession: A27811, A27812; J01149; C32668; S31103
 R:Quintrell, N.; Lebo, R.; Varmus, H.; Bishop, J.M.; Pettenati, M.J.; Le Beau, M.M.; Dia Mol. Cell. Biol. 7, 2267-2275, 1987

A:Title: Identification of a human gene (HCK) that encodes a protein-tyrosine kinase and A:Reference number: A27811; MUID:87257942; PMID:3496523
 A:Accession: A27811

A:Molecule type: mRNA

A:Residues: 1-505 <QUI>

A:Cross-references: GB:M16591

A:Note: the codon given for 3-Cys (TGC) is inconsistent with the authors' translation R:Diegeler, S.F.; Matth, J.D.; Lewis, D.B.; Perlmutter, R.M.

Mol. Cell. Biol. 7, 2276-2285, 1987

A:Title: Novel protein-tyrosine kinase gene (hck) preferentially expressed in cells of h A:Reference number: A27812; MUID:87257943; PMID:3453117

A:Accession: A27812

A:Molecule type: mRNA

A:Residues: 1-505 <ZIB>

A:Cross-references: GB:M16592; NID:G183913; PIDN:AAAS2644.1; PID:G306833

R:Radetzky, D.; Strebhardt, K.; Ruebsamen-Waigmann, H.

Gene 113, 275-280, 1992

A:Title: The genomic locus of the human hemopoietic-specific cell protein tyrosine kinase A:Reference number: J01149; MUID:92241680; PMID:1572549

A:Accession: J01149

A:Molecule type: DNA

A:Residues: 157-505 <HRA>

A:Cross-references: EMBL:X59741

R:Baranen, J.; Maekelae, T.P.; Allitalo, R.; Leivaeslahti, H.; Allitalo, K.

Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990

A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.

A:Reference number: A38268; MUID:91062389; PMID:2247464

A:Accession: C38268

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 362-417 <PAR>

C:Genetics:

A:Gene: GDB:HCK

A:Cross-references: GDB:119303; OMIM:142370

A:Map position: 20q11-20q12

A:Intons: 207/1; 258/1; 318/1; 343/3; 395/1; 439/1

C:Function:

A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho F:2-505/Product: protein-tyrosine kinase hck #status predicted <MAT>

F:64-112/Domain: SH3 homology <SH3>
 F:123-220/Domain: SH2 homology <SH2>
 F:239-497/Domain: protein kinase homology <KIN>
 F:247-255/Region: protein kinase ATP-binding motif
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F:3/Binding site: palmitate (Cys) (covalent) #status predicted
 F:269/Active site: lys #status predicted
 F:390/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 61.3%; Score 38; DB 1; Length 505;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
 DB 101 TRKEGYIPSNV 112

RESULT 8

B82379

conserved hypothetical protein VCA1085 [imported] - Vibrio cholerae (strain N16961 sero

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: B82379

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.T.; Dodson, R.J.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, I

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: B82379

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-556 <HEI>

A:Cross-references: GB:AE004434; GB:AE003853; NID:G9658531; PIDN:AAF96978.1; GSPDB:GN00:

C:Genetics:

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

A:Gene: VCA1085

A:Map position: 2

Query Match 61.3%; Score 38; DB 2; Length 556;

Best Local Similarity 58.3%; Pred. No. 36;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RSSGSIASNYVQ 13
 DB 343 KSGSIRANYAQ 354

RESULT 9

T27072
 hypothetical protein Y51A2D.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T27072

R:McMurray, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20307

A:Accession: T27072

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-606 <WIL>

A:Cross-references: EMBL:AL021497; PIDN:CAA16400.1; GSPDB:GN00023; CESP:Y51A2D.4

A:Experimental source: clone Y51A2D

C:Genetics:

A:Gene: CESP:Y51A2D.4

A:Map position: 5

A:Intons: 4/3; 76/1; 144/1; 226/3; 400/2; 446/1; 562/3

C:Superfamily: glucose transport protein

Query Match 61.3%; Score 38; DB 2; Length 606;
 Best Local Similarity 63.6%; Pred. No. 40;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 11
|:|:|:|:|:
Db 442 TRINSIANNYLQ 452

RESULT 10
A90771
hypothetical protein ECe1137 [imported] - Escherichia coli (strain O157:H7, substrain R1
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: A90771
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A/Reference number: A9629; UID:21156231; PMID:11258196
A/Accession: A90771
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-726 <NAV>
A/Cross-references: GB:BA000007; PIDN:BA034560.1; PID:g13360597; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain R1MD 0509952
C/Genetics:
A/Gene: ECe1137

Query Match 61.3%; Score 38; DB 2; Length 726;
Best Local Similarity 61.5%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|:|:|:|:|:
Db 243 TRINSIANNYLQ 255

RESULT 11
E85633
hypothetical protein yccc [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: E85633
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; UID:21074935; PMID:11206551
A/Accession: E85633
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-726 <STO>
A/Cross-references: GB:AE005174; NID:g12514246; PIDN:AA055529.1; GSPDB:GN00145; UWGP:Z13
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: yccc

Query Match 61.3%; Score 38; DB 2; Length 726;
Best Local Similarity 61.5%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|:|:|:|:|:
Db 243 TRINSIANNYLQ 255

RESULT 12
C64839
probable ATPase yccc - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C/Accession: C64839
R/Batliner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; UID:97426617; PMID:9278503
A/Accession: C64839
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-726 <BLAT>
A/Cross-references: GB:AE000200; GB:U00096; NID:g2367111; PIDN:AA04066.1; PID:g1787216,
A/Experimental source: strain K-12, substrain MG1655
C/Genetics:
A/Gene: yccc

C/Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F/4-50/Domain: transmembrane #status predicted <TM>
F/428-444/Domain: transmembrane #status predicted <TM>
F/539-546/Region: nucleotide-binding motif A (P-loop)

Query Match 61.3%; Score 38; DB 2; Length 726;
Best Local Similarity 61.5%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|:|:|:|:|:
Db 243 TRINSIANNYLQ 255

RESULT 13
PS0073
Ig kappa chain V region (38C13.V2) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jan-2000
C/Accession: PS0073
R/Levy, S.; Campbell, M.J.; Levy, R.
J. Exp. Med. 170, 1-13, 1989
A/Title: Functional immunoglobulin light chain genes are replaced by ongoing rearrangement
A/Reference number: A92781; UID:89310348; PMID:2501443
A/Accession: PS0073
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-108 <LEV>
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotrimer; immunoglobulin
F/16-91/Domain: immunoglobulin homology <IMV>

Query Match 59.7%; Score 37; DB 2; Length 108;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RSSGSIASNYV 12
|:|:|:|:|:
Db 24 RASSTISSNYL 34

RESULT 14
E59098
hypothetical protein pXOI-61 - Bacillus anthracis virulence plasmid pXOI
C/Species: Bacillus anthracis
C/Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 16-Feb-2001
C/Accession: E59098
R/Okinkaka, R.T.; Cloud, K.; Hampton, O.; Hofmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
J. Bacteriol. 181, 6509-6515, 1999
A/Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbor
A/Reference number: A59091; UID:99445483; PMID:10515943
A/Accession: E59098
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-236 <OKI>
A/Cross-references: GB:AF065404; NID:g4894216; PIDN:AA02365.1; PID:g4894277
A/Experimental source: strain Sterne

C/Genetics:
A/Gene: pXOI-61
A/Genome: plasmid
C/Superfamily: hypothetical protein pXOI-61
Query Match 59.7%; Score 37; DB 2; Length 236;

Best Local Similarity 50.0%; Pred. No. 24;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 RSSGSIASNYVQ 13
: |||: |||:
Db 25 KDSGSLITNYVE 36

RESULT 15

T30578
myosin IC - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #ext_change 08-Sep-2000
C:Accession: T30578
R:Peterson, M.D.; Novak, K.D.; Reedy, M.C.; Ruman, J.I.; Titus, M.A.
J Cell Sci. 108, 1093-1103, 1995
A:Title: Molecular genetic analysis of myoc, a Dictyostelium myosin I.
A:Reference number: Z20872; MUID:95348228; PMID:7622596
A:Accession: T30578
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1181 <PRT>
A:Cross-References: EMBL:L35323; NID:9532123; PID:9532124; PIDN:AAC37427.1
C:Genetics:
A:Gene: myoc
A:Introns: 307/1
C:Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homolo
F:18-685/Domain: myosin motor domain homology <MMO>

Query Match 59.7%; Score 37; DB 2; Length 1181;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
: |||: |||:
Db 1167 TKQIGMLPSNYVQ 1179

Search completed: January 12, 2004, 07:03:04
Job time : 14.7812 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2004, 06:47:49 ; Search time 7.10938 Seconds
(without alignments)
85.992 Million cell updates/sec

Title: US-09-829-495-64

Perfect score: 62
Sequence: 1 TRSSGSIASNYVQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	83.9	112	1	LVB6_HUMAN
2	52	83.9	131	1	LVB6_HUMAN
3	51	82.3	111	1	LVB6_HUMAN
4	43	69.4	111	1	LVB6_HUMAN
5	38	61.3	112	1	LVB6_HUMAN
6	38	61.3	504	1	HCK_MACFA
7	38	61.3	526	1	HCK_HUMAN
8	38	61.3	726	1	ETK_ECO27
9	38	61.3	726	1	ETK_ECO57
10	38	61.3	726	1	ETK_ECOLI
11	38	61.3	1399	1	Z291_HUMAN
12	37	59.7	1181	1	MYSC_DICDI
13	36	58.1	308	1	OM40_ARATH
14	36	58.1	352	1	GAL7_CAEEL
15	36	58.1	416	1	PXN1_XENLA
16	36	58.1	549	1	CET1_YEAST
17	35	56.5	278	1	PMPI_SCHPO
18	35	56.5	386	1	CMG2_HUMAN
19	35	56.5	402	1	ARGO_METTH
20	35	56.5	503	1	HCK_RAT
21	35	56.5	511	1	LYN_HUMAN
22	35	56.5	524	1	HCK_MOUSE
23	35	56.5	536	1	PNV_XIPHE
24	35	56.5	541	1	YES_MOUSE
25	35	56.5	574	1	YB2A_SCHPO
26	35	56.5	1007	1	Y741_CHLMU
27	34	54.8	37	1	CAL1_PIG
28	34	54.8	127	1	CAL2_HUMAN
29	34	54.8	294	1	TSX_ECOLI
30	34	54.8	367	1	SS81_YEAST
31	34	54.8	409	1	IVD_ARATH
32	34	54.8	740	1	GAG_SMRVH
33	34	54.8	1556	1	GLTS_SYNY3

34	33	53.2	37	1	CAL1_SHEEP	P30881 ovis aries
35	33	53.2	129	1	CAL2_HORSE	Q990T3 equus caball
36	33	53.2	178	1	FB0H_BOMMA	Q99050 bombyx mand
37	33	53.2	182	1	YBBA_ECOLI	P75866 escherichia
38	33	53.2	186	1	PRBP_BOVIN	P13696 bos taurus
39	33	53.2	186	1	PRBP_HUMAN	P30086 homo sapien
40	33	53.2	186	1	PRBP_MACFA	P48737 macaca fasc
41	33	53.2	186	1	PRBP_MOUSE	P70296 mus musculu
42	33	53.2	271	1	CRK_DROME	Q9XYM0 drosophila
43	33	53.2	325	1	Y230_ARATH	O80934 arabidopsis
44	33	53.2	336	1	ACT2_LYCCE	O96483 lycopersico
45	33	53.2	338	1	YBD1_YEAST	P38200 saccharomyc

ALIGNMENTS

RESULT 1						
LVB6_HUMAN	STANDARD;	PRT;	112 AA.			
ID	LVB6_HUMAN	STANDARD;	PRT;	112 AA.		
AC	P01722;					
DT	21-JUL-1986 (Rel. 01, Created)					
DT	21-JUL-1986 (Rel. 01, Last sequence update)					
DT	15-SEP-2003 (Rel. 42, Last annotation update)					
DE	Ig lambda chain V-VI region NIG-48.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
OX	NCBI_Taxid=9606;					
RN	[1]					
RP	SEQUENCE.					
RX	MEDLINE=80094390; PubMed=118171;					
RA	Takahashi N., Takayasu T., Isohe T., Shinoda T., Okuyama T.,					
RA	Shimizu A.;					
RT	"Comparative study on the structure of the light chains of human					
RT	immunoglobulins. II. Assignment of a new subgroup.";					
RL	J. Biochem. 86:1523-1535 (1979).					
CC	-1- MISCELLANEOUS: THIS IS A BENGE-JONES PROTEIN.					
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like domain.					
DR	PIR; A01991; L6H048.					
DR	HSSP; P01703; 7FAB.					
DR	GO; GO:0005576; C:extracellular; NAS.					
DR	GO; GO:0003823; F:antigen binding activity; NAS.					
DR	GO; GO:0006955; P:immune response; NAS.					
DR	InterPro; IPR007110; Ig-like.					
DR	InterPro; IPR003006; Ig_MHC.					
DR	InterPro; IPR003596; Ig_V.					
DR	Pfam; PF00047; Ig; 1.					
DR	SMART; SM00406; IGV; 1.					
DR	PROSITE; PS50835; IG_LIKE; 1.					
KW	Immunoglobulin V region; Bence-Jones protein.					
FT	DOMAIN 1 107					
FT	NON_TER 112					
FT	SEQUENCE 112 AA; 12152 MW; CPEB307BC527A384 CRC64;					
Query Match						
Best Local Similarity		83.9%;	Score 52;	DB 1;	Length 112;	
Matches 11; Conservative		1;	Mismatches	1;	Indels	0;
Matches 11; Conservative		1;	Mismatches	1;	Indels	0;
QY	1 TRSSGSIASNYVQ 13					
DB	23 TRSSGSIASNYVQ 35					
RESULT 2						
LVB6_HUMAN	STANDARD;	PRT;	131 AA.			
ID	LVB6_HUMAN	STANDARD;	PRT;	131 AA.		
AC	P06319;					
DT	01-JAN-1988 (Rel. 06, Created)					
DT	01-JAN-1988 (Rel. 06, Last sequence update)					
DT	15-JUL-1999 (Rel. 38, Last annotation update)					
DE	Ig lambda chain V-VI region E84 precursor.					
OS	Homo sapiens (Human).					

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1] Y
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215660; PubMed=3923440;
 RA Anderson M.L.M., Brown L., McKenzie E., Kellow J.E., Young B.D.;
 RT "Cloning and sequence analysis of an Ig lambda light chain mRNA
 expressed in the Burkitt's lymphoma cell line EB4."
 RL Nucleic Acids Res. 13:2931-2941(1985).
 DR PIR; A01900; LSHUB.
 DR HSSP; P01709; 2MCG.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; P:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KM Immunoglobulin V region; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 131 IG LAMBDA CHAIN V-VI REGION EB4.
 FT DOMAIN 20 41 FRAMEWORK-1.
 FT DOMAIN 42 54 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 55 69 FRAMEWORK-2.
 FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 77 110 FRAMEWORK-3.
 FT DOMAIN 111 118 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 119 131 FRAMEWORK-4.
 FT DISULFID 41 110 BY SIMILARITY.
 FT NON TER 131 131
 SQ SEQUENCE 131 AA; 14147 MW; 02A0179C8C05C2CD CRC64;

Query Match 83.3%; Score 52; DB 1; Length 131;
 Best Local Similarity 84.6%; Pred. No. 0.0041;
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 42 TGNSSGSIASNYVQ 54

RESULT 3
 ID LV6D_HUMAN STANDARD; PRT; 111 AA.
 AC P06318;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig lambda chain V-VI region WLT.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86122667; PubMed=4089539;
 RA Dwalet F.E., Strako K., Benson M.D.;
 RT "Amino acid sequence of a lambda VI primary (AL) amyloid protein
 (WLT)."
 RL Scand. J. Immunol. 22:653-660(1985).
 DR PIR; A01989; LSHUT.
 DR HSSP; P01709; 2MCG.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; P:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR SMART; SM00406; IGV; 1.

DR PROSITE; PS50835; IG_LIKE; 1.
 KM Immunoglobulin V region.
 FT DOMAIN 1 22 FRAMEWORK-1.
 FT DOMAIN 23 35 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 36 50 FRAMEWORK-2.
 FT DOMAIN 51 57 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 58 91 FRAMEWORK-3.
 FT DOMAIN 92 101 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 102 111 FRAMEWORK-4.
 FT DISULFID 22 91 BY SIMILARITY.
 FT NON TER 111 111
 SQ SEQUENCE 111 AA; 11966 MW; 0C88B2FE37BC24F CRC64;

Query Match 82.3%; Score 51; DB 1; Length 111;
 Best Local Similarity 84.6%; Pred. No. 0.0054;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 23 TGNSSGSIASNYVQ 35

RESULT 4
 ID LV6C_HUMAN STANDARD; PRT; 111 AA.
 AC P06317;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig lambda chain V-VI region STU.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RA Solomon A., Kyle R.A., Frangione B.;
 RT "Light chain variable region subgroups of monoclonal immunoglobulins
 in amyloidosis AL.";
 RL (in) Glenner G.G., Osseman E.F., Benditt E.P., Calkins E.,
 Cohen A.S., Zucker-Franklin D. (eds.); Plenum Press, New York (1986).
 RL Amyloidosis, pp.449-462, Plenum Press, New York (1986).
 DR PIR; A01988; LSHUT.
 DR PDB; 1CD0; 06-MAR-00.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KM Immunoglobulin V region; 3D-structure.
 FT DOMAIN 1 22 FRAMEWORK-1.
 FT DOMAIN 23 35 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 36 50 FRAMEWORK-2.
 FT DOMAIN 51 57 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 58 91 FRAMEWORK-3.
 FT DOMAIN 92 100 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 101 111 FRAMEWORK-4.
 FT DISULFID 22 91 BY SIMILARITY.
 FT NON TER 111 111
 SQ SEQUENCE 111 AA; 12247 MW; 0941DD547D983598 CRC64;

Query Match 69.4%; Score 43; DB 1; Length 111;
 Best Local Similarity 69.2%; Pred. No. 0.2;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 23 TRSDGTINGYVQ 35

RESULT 5
 ID LV6A_HUMAN

ID	LVFA HUMAN	STANDARD;	PRT;	112 AA.
AC	P01721;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ig lambda chain V-VI region AR.			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
CC	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE (AMYLLOID PROTEIN AR).			
RA	MEDLINE=82091000; PubMed=6797401;			
RA	Sletten K., Natvig J.B., Husby G., Junl J.;			
RT	"the complete amino acid sequence of a prototype			
RT	immunoglobulin-lambda light-chain-type amyloid-fibril protein AR."			
RL	Biochem. J. 195;561-572(1981).			
CC	-I- MISCELLANEOUS: ABOUT HALF OF THE LAMBDA CHAIN C REGION IS MISSING FROM THIS PROTEIN.			
CC	-I- MISCELLANEOUS: THIS PROTEIN WAS ISOLATED FROM THE SPLEEN OF A PATIENT WITH AMYLOIDOSIS.			
CC	-I- SIMILARITY: Contains 1 immunoglobulin-like domain.			
DR	PIR; A01987; L6HUR.			
DR	HSSP; P01709; 2MCG.			
DR	GO; GO:0005576; C:extracellular; NAS.			
DR	GO; GO:0003823; F:antigen binding activity; NAS.			
DR	GO; GO:0006955; P:immune response; NAS.			
DR	InterPro; IPR007110; IG-1Like.			
DR	InterPro; IPR003006; IG_MHC.			
DR	InterPro; IPR003596; IG_V.			
DR	Pfam; PF00047; Ig_1.			
DR	SMART; SM00406; IGV_1.			
DR	PROSITE; PS50835; IG_LIKE; 1.			
KW	Immunoglobulin V region; Amyloid.			
DR	DOMAIN 1 107			
FT	NON TER 112 112			
FT	SEQUENCE 112 AA; 11918 MW; 570BCD9A368EF1FE CRC64;			
QY	Query Match	61.3%;	Score 38;	DB 1; Length 112;
Db	Best Local Similarity	61.5%;	Pred. No. 2;	
	Matches 8; Conservative	2;	Mismatches 3;	Indels 0; Gaps 0;
	1 TRSSGSISANYQ 13			
	23 TGSQGSISDSFVQ 35			
RESULT 6				
HCK	MACFA	STANDARD;	PRT;	504 AA.
AC	Q95M30;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Tyrosine-protein kinase HCK (EC 2.7.1.112) (p56-HCK) (Hemopoietic cell kinase).			
GN	HCK.			
OS	Macaca fascicularis (Crap eating macaque) (Cynomolgus monkey).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;			
CC	Cercopitheciae; Macaca.			
CC	NCBI_TaxID=9541;			
CC	(1)			
CC	SEQUENCE FROM N.A.			
RA	Picard C.;			
RL	Thesis (2001), University of Marseille, France.			
CC	-I- FUNCTION: MAY SERVE AS PART OF A SIGNALING PATHWAY COUPLING THE FC RECEPTOR TO THE ACTIVATION OF THE RESPIRATORY BURST. MAY ALSO CONTRIBUTE TO NEUTROPHIL MIGRATION AND MAY REGULATE THE DEGRANULATION PROCESS OF NEUTROPHILS (By similarity).			
CC	-I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.			
CC	-I- SUBCELLULAR LOCATION: Membrane-associated (By similarity).			

```

CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/1998-01-01).
CC or send an email to license@ebi.ac.uk.
CC -----
CC EMBL, AJ320181, CAC44031.1; --
CC InterPro, IPR000719, Prot_Kinase.
CC InterPro, IPR000980, SH2.
CC InterPro, IPR001452, SH3.
CC InterPro, IPR001245, Tyr_PKinase.
CC Pfam, PF00069, pkinase, 1.
CC Pfam, PF00017, SH2, 1.
CC Pfam, PF00018, SH3, 1.
CC PRINTS, PRO0401, SH2DOMAIN.
CC PRINTS, PRO0452, SH3DOMAIN.
CC PRINTS, PRO0109, TYRKINASE.
CC ProDom, PD000001, Prot_kinase, 1.
CC ProDom, PD000093, SH2, 1.
CC ProDom, PD000066, SH3, 1.
CC SMART, SM00252, SH2, 1.
CC SMART, SM00326, SH3, 1.
CC SMART, SM00219, Tyrc, 1.
CC PROSITE, PS00107, PROTEIN_KINASE_ATP, 1.
CC PROSITE, PS00109, PROTEIN_KINASE_TYR, 1.
CC PROSITE, PS50011, PROTEIN_KINASE_DOM, 1.
CC PROSITE, PS50002, SH2, 1.
CC PROSITE, PS50002, SH3, 1.
CC Transferrase: Tyrosine-protein kinase: Phosphorylation; ATP-binding;
KW Myristate; SH2 domain; SH3 domain.
KW DOMAIN
FT 122 219 SH3.
FT DOMAIN
FT 122 219 SH2.
FT NP_BIND
FT 246 254 ATP (BY SIMILARITY).
FT BINDING
FT 268 268 ATP (BY SIMILARITY).
FT ACT_SITE
FT 359 359 BY SIMILARITY.
FT LIPID
FT 2 2 MYRISTATE (BY SIMILARITY).
FT MOD_RES
FT 389 389 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 504 AA; 57096 MW; 53B29322D2DE3423 CRC64;

Query Match 61.3%; Score 38; DB 1; Length 504;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TRSGSGIASNVY 12
Db 100 TRKEGYIPSNV 111

RESULT 7
HCK_HUMAN STANDARD; PRT; 526 AA.
AC P06631; Q96CC0; Q9H5Y5; Q9NUA4; Q9UMJ5;
DT 01-AUG-1988 (Rel. 08, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase HCK (EC 2.7.1.112) (p59-HCK/p60-HCK)
DE (Hemopoietic cell kinase).
DE HCK.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 22-526 FROM N.A.
RX MEDLINE=87257942; Pubmed=3496523;

```

RA Quintrell N., Lebo R., Varmus H., Bishop J.M., Pettenati M.J.,
 RA le Beau M.M., Diaz M.O., Rowley J.D.;
 RT "Identification of a human gene (HCK) that encodes a protein-tyrosine
 RT kinase and is expressed in hemopoietic cells.";
 RN Mol. Cell. Biol. 7:2267-2275(1987).
 RN [12]
 RP SEQUENCE OF 22-526 FROM N.A.
 RX MEDLINE=87257943; PubMed=3453117;
 RA Ziegler S.F., Marsh J.D., Lewis D.B., Perlmuter R.M.;
 RT "Novel protein-tyrosine kinase gene (nck) preferentially expressed in
 RT cells of hematopoietic origin";
 RL Mol. Cell. Biol. 7:2276-2285(1987).
 RN [3]
 RP SEQUENCE OF 22-526 FROM N.A.
 RX TISSUE=B-cell;
 RA MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko J., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywiński M.I., Skalska U., Smalhus D.E.,
 RA Scherach A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 22-526 FROM N.A.
 RX TISSUE=leal mucosa;
 RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
 RA Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y., Ota T., Suzuki Y.,
 RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
 RA Isogai T., Sugano S.;
 RT "NEO human cDNA sequencing project";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton A.J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths M.N.D., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G., Lawlor S.,
 RA Leharvalaino M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Symons N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [6]
 RP SEQUENCE OF 179-526 FROM N.A.
 RX TISSUE=spleen;
 RA MEDLINE=92241680; PubMed=1572549;
 RA Hradetzky D., Strebhardt K., Ruesamen-Waigmann H.;
 RT "The genomic locus of the human hemopoietic-specific cell protein
 RT tyrosine kinase (PTK)-encoding gene (HCK) confirms conservation of
 RT exon-intron structure among human PTKs of the src family.";
 RL Gene 113:215-280(1992).
 RN [7]
 RP SEQUENCE OF 1-22 FROM N.A., AND ALTERNATIVE INITIATION.
 RX MEDLINE=91342636; PubMed=1875927;
 RA Lock P., Ralph S., Stanley E., Boulter I., Ramsay R., Dunn A.R.;
 RT "Two isoforms of murine hck, generated by utilization of alternative
 RT translational initiation codons, exhibit different patterns of
 RT subcellular localization";
 RL Mol. Cell. Biol. 11:4363-4370(1991).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 78-526.
 RX MEDLINE=97177106; PubMed=9024658;
 RA Sichert F., Moarefi I., Kuriyan J.;
 RT "Crystal structure of the Src family tyrosine kinase Hck";
 RL Nature 385:602-609(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 81-137.
 RX MEDLINE=98453315; PubMed=9778343;
 RA Arold S., O'Brien R., Franken P., Strub M.P., Hoh F., Dumas C.,
 RA Ladbury J.E.;
 RT "RT loop flexibility enhances the specificity of Src family SH3
 RT domains for HIV-1 Nef";
 RL Biochemistry 37:14683-14691(1998).
 RN [10]
 RP STRUCTURE BY NMR OF 78-138.
 RX MEDLINE=98239731; PubMed=9571048;
 RA Horita D.A., Baldisseri D.M., Zhang W., Altieri A.S., Smithgall T.E.,
 RA Gmeiner W.H., Byrd R.A.;
 RT "Solution structure of the human Hck SH3 domain and identification of
 RT its ligand binding site";
 RL J. Mol. Biol. 278:253-265(1998).
 RN [11]
 RP STRUCTURE BY NMR OF 139-245.
 RX MEDLINE=97263487; PubMed=9109402;
 RA Zhang W., Smithgall T.E., Gmeiner W.H.;
 RT "Sequential assignment and secondary structure determination for the
 RT Src homology 2 domain of hematopoietic cellular kinase";
 RL FEBS Lett. 406:131-135(1997).
 RN [12]
 RP FUNCTION: MAY SERVE AS PART OF A SIGNALING PATHWAY COUPLING THE FC
 CC CONTRIBUTOR TO NEUTROPHIL MIGRATION AND MAY REGULATE THE
 CC DEGRANULATION PROCESS OF NEUTROPHILS.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Isoform p60-HCK and isoform p59-HCK are
 CC associated with membranes. Isoform p60-HCK is also cytoplasmic (By
 CC similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative Initiation;
 CC Comment2: Isoforms, p60-HCK (shown here) and p59-HCK, are
 CC produced by alternative initiation;
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CELLS OF THE
 CC MYELOID AND B-LYMPHOID LINEAGES.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
 CC SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL, M16591; AAA52643.1; -
DR EMBL, M16592; AAA52644.1; -
DR EMBL, BC014435; AAH14435.1; -
DR EMBL, AK026432; BAB15482.1; -
DR EMBL, AL049539; CAB75606.1; -
DR EMBL, X58741; CAA1565.2; -
DR EMBL, X58742; CAA1565.2; JOINED.
DR EMBL, X58743; CAA1565.2; JOINED.
DR PIR, A27811; TVDHUC.
DR PDB, 3HCK; 20-AUG-97.
DR PDB, 3HCK; 15-OCT-97.
DR PDB, 4HCK; 17-JUN-98.
DR PDB, 5HCK; 17-JUN-98.
DR PDB, 1AD5; 15-MAY-97.
DR PDB, 1BU1; 11-NOV-98.
DR PDB, 1OCF; 08-JUN-99.
DR Genew; HGNC:4640; HCK.
DR MIM; 142370; -
DR GO; GO:0004713; F-protein tyrosine kinase activity; TAS.
DR GO; GO:0007498; P-mesoderm development; TAS.
DR GO; GO:0006468; P-protein amino acid phosphorylation; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00252; SH2; 1.

Query Match 61.3%; Score 38; DB 1; Length 526;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 12
DB 122 TRKEGIPSNV 133

RESULT 8
ETK_ECO27 STANDARD; PRT; 726 AA.
ID ETK_ECO27 STANDARD; PRT; 726 AA.
AC P58764;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase etk (EC 2.7.1.112).
GN ETK.
OS *Escherichia coli* O127:H6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=168807;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 1-17, AND CHARACTERIZATION.
RC STRAIN=O127:H6 / E2348/69;
RC MEDLINE=92928177; PubMed=10369665;
RA Ilan O.A., Bloch Y., Frankel G., Ulirich H., Geider K., Rosenshine I.,
RT "Protein tyrosine kinases in bacterial pathogens are associated with
virulence and production of exopolysaccharide";
RL EMBL J. 183241-3248(1999).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Inner membrane.

CC -1- PTM: Autophosphorylated. Dephosphorylated by etp (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ETK/WZC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL, AJ238695; CAB43868.1; -
DR InterPro; IPR003856; LPS_WZC_MPA.
DR Pfam; PF02706; wzz; 1.
KW Transferase; Tyrosine-protein kinase; Phosphorylation; Transmembrane;
KW Inner membrane.
FT TRANSMEM 33
FT TRANSMEM 425 445 POTENTIAL.
SQ SEQUENCE 726 AA; 81082 MW; 91F97DF8AD8BE302 CRC64;

Query Match 61.3%; Score 38; DB 1; Length 726;
Best Local Similarity 61.5%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
DB 243 TRILNSIANNYIQ 255

RESULT 9
ETK_ECO57 STANDARD; PRT; 726 AA.
ID ETK_ECO57 STANDARD; PRT; 726 AA.
AC O8XC28;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase etk (EC 2.7.1.112).
GN ETK OR Z1398 OR ECS1137.
OS *Escherichia coli* O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimailanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.,
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7,"
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11238796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
O157:H7 and genomic comparison with a laboratory strain K-12,"
RL DNA Res. 8:11-22(2001).
RN [1]
RP -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC -1- PTM: Autophosphorylated. Dephosphorylated by etp (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ETK/WZC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE005292; AAG55529.1; -
 DR EMBL; AP002554; BAB34560.1; -
 DR PIR; A90771; A90771.
 DR PIR; E85633; E85633.
 DR InterPro; IPR003856; LPS_Wzz_MPA.
 DR Pfam; PF02706; wzz; 1.
 DR Transferrase; Tyrosine-protein kinase; Phosphorylation; Transmembrane;
 KM Inner membrane; Complete proteome.
 FT TRANSMEM 33
 FT TRANSF 45
 FT SEQUENCE 726 AA; 81259 MW; 7BEA191209711BDC CRC64;
 SQ

Query Match 61.3%; Score 38; DB 1; Length 726;
 Best Local Similarity 61.5%; Pred. No. 15;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 243 TRILINSIANNYLQ 255

RESULT 10
 ID ETK_ECOLI STANDARD; PRT; 726 AA.
 AC P38134; P75879;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase *etk* (EC 2.7.1.112).
 ETK OR B0981.
 OS Escherichia coli.
 ON Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=562;
 RN
 RP SEQUENCE FROM N.A.
 RX STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.D.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishino K., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takenoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horituchi T.;
 RT "A 7.8-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 RN
 RP SEQUENCE OF 628-726 FROM N.A.
 RX MEDLINE=90368616; PubMed=2168385;
 RA Dassa J., Marck C., Bognet P.L.;
 RT "The complete nucleotide sequence of the Escherichia coli gene *appa*
 RT glucose-1-phosphatase.";
 RL J. Bacteriol. 172:5497-5500(1990).
 RN
 RP IDENTIFICATION.

RX MEDLINE=95075659; PubMed=7984428;
 RA Borodovsky M., Rudd K.E., Koonin E.V.;
 RT "Intrinsic and extrinsic approaches for detecting genes in a
 RT bacterial genome.";
 RL Nucleic Acids Res. 22:4756-4767(1994).
 RN
 RP CHARACTERIZATION.
 RC STRAIN=K12 / JM109;
 RX MEDLINE=20545593; PubMed=11090276;
 RA Vincent C., Ducloux B., Grangeasse C., Vaganay E., Riberty M.,
 RA Cozzone A.J., Doublet P.;
 RT "Relationship between exopolysaccharide production and protein-
 RT tyrosine phosphorylation in gram-negative bacteria.";
 RL J. Mol. Biol. 304:311-321(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Inner membrane (By similarity).
 CC -1- PTM: Autophosphorylated. Dephosphorylated by *etp*.
 CC -1- SIMILARITY: BELONGS TO THE ETK/WZZ FAMILY.
 CC -1- CAUTION: Seems to be expressed only in enteropathogenic E.coli
 CC strains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE000200; AAC74066.1; -
 DR EMBL; D90735; BAA35746.1; -
 DR EMBL; M58708; -; NOT_ANNOTATED_CDS.
 DR PIR; C64839; C64839.
 DR PhosSite; P38134; etk.
 DR EcoGene; EG1826; etk.
 DR InterPro; IPR003856; LPS_Wzz_MPA.
 DR Pfam; PF02706; wzz; 1.
 KM Transferrase; Tyrosine-protein kinase; Phosphorylation; Transmembrane;
 KM Inner membrane; Complete proteome.
 FT TRANSMEM 33
 FT TRANSF 45
 FT TRANSMEM 425
 FT TRANSF 445
 FT CONFLICT 726
 FT SEQUENCE 726 AA; 81241 MW; BAA060F59680DA22 CRC64;
 SQ

Query Match 61.3%; Score 38; DB 1; Length 726;
 Best Local Similarity 61.5%; Pred. No. 15;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
 DB 243 TRILINSIANNYLQ 255

RESULT 11
 ID Z291 HUMAN STANDARD; PRT; 1399 AA.
 AC Q9BY12;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 291.
 GN ZNF291.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RA Carim L., Sumoy L., Estivill X., Becarceller M.;
 RT "Identification and characterization of ZNF291, a novel protein on
 RT chromosome 15q24.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF242528; AAK29205.1; -
DR Genew; HGNC:13081; ZNF291.
DR InterPro; IPR007087; Znf_C2H2.
DR SMART; SM00355; Znf_C2H2_1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; FALSE_NEG.
KW Zinc-finger; Metal-binding.
FT DOMAIN 394 769 GLU-RICH.
FT ZN_FING 791 815 C2H2-TYPE.
SQ SEQUENCE 1399 AA; 158156 MW; 905C63CD908B5013 CRC64;

Query Match 61.3%; Score 38; DB 1; Length 1399;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RSSGSIASNYVQ 13
Db 354 KNSGSIRDNVVR 365
:::|||||
ID MISC DICDI STANDARD; PRT; 1181 AA.
AC P42532;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin IC heavy chain.
GN MYOC OR DMIC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RX MEDLINE=95348228; PubMed=7622596;
RA Peterson M.D., Novak K.D., Reedy M.C., Ruman J.I., Titus M.A.;
RT "Molecular genetic analysis of myoc, a Dictyostelium myosin I.";
RL J. Cell Sci. 108:1093-1103(1995).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC -1- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY
CC AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L35323; AAC37427.1; -
DR FIR; T30578; T30578.
DR HSSP; P08799; 1MND.
DR DictyDb; DD01090; myoc.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR000108; Neu_cyt_fact_2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00063; myosin_head; 2.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00193; MYOSINHEAVY.

DR PRINTS; PR00499; P67PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000355; myosin_head; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00242; MYSC; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50022; SH3; 1.
KW Myosin; Actin-binding; ATP-binding; Chemotaxis; SH3 domain;
KW Multigene family.
FT DOMAIN 1 2 MYOSIN HEAD-LIKE.
FT DOMAIN 2 1181 NON ALPHA-HELICAL, C-TERMINAL DOMAIN.
FT NP_BIND 109 116 ATP (POTENTIAL).
FT DOMAIN 1122 1181 SH3.
SQ SEQUENCE 1181 AA; 132915 MW; 5EB1E47F0CA8803 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 1181;
Best Local Similarity 53.8%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
Db 1167 TKQIGMLPSNYVQ 1179
:::|||||
ID OM40 ARATH STANDARD; PRT; 308 AA.
AC 09LH5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable mitochondrial import receptor subunit TOM40 homolog
DE (Translocase of outer membrane 40 kDa subunit homolog).
GN TOM40 OR AR3G20000 OR MZE19.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. TAC
RT Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
RL DNA Res. 7:217-221(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-15; 21-42; 63-80; 207-219 AND 245-258.
RC STRAIN=cv. Columbia;
RX PubMed=11161051;
RA Werhahn W., Niemeyer A., Jaensch L., Kruff V., Schmitz U.K.,
RA Braun H.-P.;
RT "Purification and characterization of the preproteins translocase of
RT the outer mitochondrial membrane from Arabidopsis thaliana.
RT Identification of multiple forms of TOM20.";
RL Plant Physiol. 125:943-954(2001).
RN [4]
RP SEQUENCE OF 1-15 AND 246-258, AND SUBCELLULAR LOCATION.
RC TISSUE=leaf, and stem;
RX MEDLINE=21608403; PubMed=11743114;
RA Kruff V., Eubel H., Jaensch L., Werhahn W., Braun H.-P.;
RT "Proteomic approach to identify novel mitochondrial proteins in
RT Arabidopsis.";
RL Plant Physiol. 127:1694-1710(2001).

CC -1- FUNCTION: CENTRAL COMPONENT OF THE RECEPTOR COMPLEX RESPONSIBLE
CC FOR THE RECOGNITION AND TRANSLOCATION OF CYTOSOLICALLY SYNTHESIZED
CC MITOCHONDRIAL PREPROTEINS. TOGETHER WITH TOM22 FUNCTIONS AS THE
CC TRANSIT PEPTIDE RECEPTOR AT THE SURFACE OF THE MITOCHONDRION OUTER
CC MEMBRANE AND FACILITATES THE MOVEMENT OF PREPROTEINS INTO THE
CC TRANSLOCATION PORE. DIRECTLY INVOLVED IN THE PORE FORMATION.
CC -1- SUBUNIT: FORMS PART OF MITOCHONDRIAL RECEPTOR COMPLEX.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC outer membrane.
CC -1- SIMILARITY: BELONGS TO THE TOM40 FAMILY.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP002050; BAB03165.1; ALT_SEQ.
CC EMBL: AY050925; AAK93602.1; -.
CC DR InterPro: IPR001925; Forin_Euk.
CC DR Pfam: PF01459; Euk_porin; 1.
CC KM Transport: Protein transport; Outer membrane; Mitochondrion;
CC Transmembrane.
CC FT INIT MET 0 0 INTERMEMBRANE (POTENTIAL).
CC FT DOMAIN 1 70 POTENTIAL.
CC FT TRANSLEM 71 86 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 87 187
CC SQ SEQUENCE 308 AA; 34119 MW; A013F8798BAAAT720 CRC64;
CC
CC Query Match 58.1%; Score 36; DB 1; Length 308;
CC Best Local Similarity 54.5%; Pred. No. 15;
CC Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 3 SSGSIANSNYV 13
CC Db 216 STGAVNNYV 226
CC
CC RESULT 14
CC GALT_CABEL
CC ID GALT_CABEL STANDARD; PRT; 352 AA.
CC AC Q27535;
CC DT 15-SEP-2003 (Rel. 42, Created)
CC DT 15-SEP-2003 (Rel. 42, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Probable galactose-1-phosphate uridylyltransferase (EC 2.7.7.12)
CC DE (Gal-1-P uridylyltransferase) (UDP-glucose--hexose-1-phosphate
CC uridylyltransferase).
CC DE ZK1056.3.
CC OS Caenorhabditis elegans.
CC OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
CC OC Rhabditidae; Peldoderinae; Caenorhabditis.
CC OX NCBI_TaxID=6239;
CC OX [1]
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Bristol N2;
CC RA Mortimore B.J.;
CC RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC RL -1- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
CC alpha-D-glucose 1-phosphate + UDP-galactose.
CC -1- COFACTOR: Binds 1 zinc and 1 iron ion per subunit (Potential).
CC -1- PATHWAY: Galactose metabolism; second step.
CC -1- SIMILARITY: Belongs to the galactose-1-phosphate
CC uridylyltransferase family 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z35604; CAAB4679.1; -.
CC DR PIR: T27677; T27677.
CC DR HSSP: P09148; IHXO.
CC DR WormPep: ZK1058.3; CE01103.
CC DR InterPro: IPR001937; Galp_UDPtransf1.
CC DR InterPro: IPR005851; Galp_Utransf_1.
CC DR InterPro: IPR005850; Galp_Utransf_C.
CC DR InterPro: IPR005849; Galp_Utransf_N.
CC DR Pfam: PF01087; Galp_UDP_transf; 1.
CC DR Pfam: PF02744; Galp_UDP_tr_C; 1.
CC DR ProDom: PD005051; Galp_UDPtransf1; 1.
CC DR TIGRPFAM: TIGR00209; gal1; 1.
CC DR PROSITE: PS00117; Gal_P_UDP_TRANSF_1; 1.
CC KM Transferase; Nucleotidyltransferase; Galactose metabolism; Zinc; Iron;
CC Metal-binding.
CC FT ACT SITE 167 167 NUCLEOPHILE (BY SIMILARITY).
CC FT METAL 165 165 ZINC (POTENTIAL).
CC FT METAL 300 300 IRON (POTENTIAL).
CC SQ SEQUENCE 352 AA; 41042 MW; F1B433A926C47C4C CRC64;
CC
CC Query Match 58.1%; Score 36; DB 1; Length 352;
CC Best Local Similarity 66.7%; Pred. No. 17;
CC Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 1 TRSSGSIANSNYV 12
CC Db 59 TRSSGSIANSNYV 70
CC
CC RESULT 15
CC PXN1_XENIA
CC ID PXN1_XENIA STANDARD; PRT; 416 AA.
CC AC P49263;
CC DT 01-FEB-1996 (Rel. 33, Created)
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
CC DE Pentraxin fusion protein precursor.
CC DE PXN1.
CC OS Xenopus laevis (African clawed frog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
CC OC Xenopodinae; Xenopus.
CC OX NCBI_TaxID=8355;
CC OX [1]
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Liver;
CC RX MEDLINE=94052237; PubMed=7694301;
CC RA Seery L.T., Schoenberg D.R., Barbaux S., Sharp P.M., Whitehead A.S.;
CC RT "Identification of a novel member of the pentraxin family in Xenopus
CC laevis".
CC RT Proc. R. Soc. Lond. B, Biol. Sci. 253:263-270(1993).
CC RL -1- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L19881; AAC38013.1; -.
CC DR HSSP: P07629; IHAS.
CC DR InterPro: IPR006585; FTPI.
CC DR InterPro: IPR001759; Pentaxin.
CC DR Pfam: PF00354; pentaxin; 1.
CC DR PRINTS: PR00895; PENTAXIN.
CC DR ProDom: PD002153; Pentaxin; 1.
CC DR SMART: SM00607; FTP; 1.

DR SMART; SM00159; PTX; 1.
 KW PROSITE; PS00289; PENTAXIN; 1.
 FT PENTAXIN; Signal.
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 15 416 PENTAXIN FUSION PROTEIN.
 FT DOMAIN 232 416 PENTAXIN.
 FT DISULFID 251 311 BY SIMILARITY.
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 416 AA; 47241 MW; 62EDE45E913A2BD7 CRC64;

Query Match 58.1%; Score 36; DB 1; Length 416;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 GSIASTYVQ 13
 ||:|||||:
 Db 64 GSIASTYME 72

Search completed: January 12, 2004, 06:59:43
 Job time : 11.1094 secs

THIS PAGE BLANK (USPTO)


```

096JD0      PRELIMINARY;      PRT;      116 AA.
ID 096JD0
AC 096JD0
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Amyloid lambda 6 light chain variable region SAR (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Perfecti V., Casarini S., Colli Vignarelli M., Merlini G.;
RT "Amyloid lambda 6 light chain variable region SAR."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF267875; AAK58587.1; -.
DR InterPro; IPR007110; IG_1ike.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_1.
DR SMART; SM00406; IGV_1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 116 AA; 12294 MW; F7B0E9F49FAE369E CRC64;

Query Match 83.9%; Score 52; DB 4; Length 116;
Best Local Similarity 84.6%; Pred. NO. 0.022;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
DB 23 TGSSGSIAINYVQ 35

RESULT 3
09V9J3      PRELIMINARY;      PRT;      517 AA.
ID 09V9J3; 094879; 018369; Q26297;
AC 09V9J3; 094879; 018369; Q26297;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Tyrosine-protein kinase Src42A (RC 2.7.1.112).
GN SRC42A OR SRC41 OR TK5 OR CG7873.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RC STRAIN=Canton-S; TISSUE=Pupae;
RX MEDLINE=96268448; PubMed=8682295;
RA Takahashi F., Endo S., Kojima T., Saigo K.;
RT "Regulation of cell-cell contacts in developing Drosophila eyes by
RT Dsrc4, a new, close relative of vertebrate c-src."
RL Genes Dev. 10:1645-1656(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

```

```

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasbo P., Lei Y., Levytsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Peltman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spierke R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
RA Wang Z.-Y., Wesserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragae V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 374-428 FROM N.A.
RX MEDLINE=92008631; PubMed=1915852;
RA Shishido E., Emori Y., Saigo K.;
RT "Identification of seven novel protein-tyrosine kinase genes of
RT Drosophila by the polymerase chain reaction."
RL FEBS Lett. 289:235-238(1991).
RN [5]
RP SEQUENCE OF 376-427 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=98401146; PubMed=9731193;
RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;
RT "Sampling the genomic pool of protein tyrosine kinase genes using the
RT polymerase chain reaction with genomic DNA."
RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
RN [6]
RP FUNCTION: ESSENTIAL FOR CORRECT EYE MORPHOGENESIS (OMMATIDIAL R7
RP NEURON FORMATION), THIS REQUIRES THE RAS1/RAF SIGNAL TRANSDUCTION
RP PATHWAY. MAY BE INVOLVED IN THE REGULATION OF CYTOSKELETON
RP ORGANIZATION AND CELL-CELL CONTACTS IN DEVELOPING OMMATIDIA.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC TYROSINE PHOSPHATE.
CC -1- TISSUE SPECIFICITY: UBQUITOUS IN EARLY EMBRYOS, IN STAGES 13-16
CC EXPRESSION IS SEEN IN VISCERAL MESODERM, HINDGUT, BRAIN, ANAL PADS
CC AND VENTRAL GANGLIONS. IN LARVAE, EXPRESSION IS IN CNS, WING DISK,
CC LEG DISK AND PHOTORECEPTOR PRECURSORS IN THE EYE-ANTENNA DISKS
CC POSTERIOR TO THE MORPHOGENETIC FIBROW
CC -1- DEVELOPMENTAL STAGE: IN EARLY EMBRYOS EXPRESSION IS VERY LOW,
CC EXPRESSION INCREASES DURING EMBRYOGENESIS. ALSO EXPRESSED IN
CC LARVAE AND PUPAE.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE SRC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; D42125; BAA07705.1; -.
DR EMBL; AF003784; AAF57295.1; -.

```

DR EMBL; AY058652; AAL13881.1; -.
DR EMBL; S55977; AAB19907.1; -.
DR EMBL; AJ002911; CA05746.1; -.
DR HSSP; P00523; ISR1.
DR FlyBase; FBgn004603; Src42A.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000066; SH3; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00002; SH3; 1.
DR TRANSFERASE; Kinase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Developmental protein.
FT DOMAIN 63 124
FT DOMAIN 130 222
FT NP_BIND 248 504
FT BINDING 254 262
FT ACT_SITE 276 276
FT CONFLICT 370 370
FT CONFLICT 65 65
FT CONFLICT 376 376
FT CONFLICT 381 386
SQ SEQUENCE 517 AA; 59069 MW; 1EF196F4D7AE61E9 CRC64;

Query Match 67.7%; Score 42; DB 5; Length 517;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TRSGSIASNYV 12
Db 108 TRSEGYPISNYV 119

RESULT 4
Q96JD2 PRELIMINARY; PRT; 112 AA.
AC Q96JD2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Amyloid lambda 6 light chain variable region NEG (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow;
RA Perfect V., Casarin S., Colla Vignarelli M., Merlini G.,
RT "Amyloid lambda 6 light chain variable region NEG";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A261873; AKS58585.1; -.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_1.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 11908 MW; 080B4B37E2360B06 CRC64;

Query Match 66.1%; Score 41; DB 4; Length 112;
Best Local Similarity 76.9%; Pred. No. 3.2;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 TRSGSIASNYVQ 13
Db 23 TGSGRISNSVQ 35

RESULT 5
ID 014244 PRELIMINARY; PRT; 778 AA.
AC 014244;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN SPAC6F6.13C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972b-;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98981; CAB11736.1; -.
DR GeneDB; SPombe; SPAC6F6.13C; -.
DR InterPro; IPR000734; Lipase.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hypothetical protein.
SQ SEQUENCE 778 AA; 86129 MW; 080B2AC38E5D1796 CRC64;

Query Match 64.5%; Score 40; DB 3; Length 778;
Best Local Similarity 72.7%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 SSGSIASNYVQ 13
Db 39 SKGSITENYVQ 49

RESULT 6
Q8XIV8 PRELIMINARY; PRT; 945 AA.
AC Q8XIV8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Exo-beta-1,3-glucanase (EC 3.2.1.58).
GN EXG8.
OS Aspergillus saitoi (Aspergillus phoenicis).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.
OX NCBI_TaxID=18681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC1432;
RA Oda K., Abe K., Kasahara S., Nakajima T.;
RT "Cloning and nucleotide sequencing of the exo-beta-1,3-glucanase gene
RT (exs) from Aspergillus saitoi";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB070739; BAB83607.1; -.
DR InterPro; IPR006162; Peptinase attach.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
KW Hydrolyase; Glycosidase.
SQ SEQUENCE 945 AA; 99332 MW; 58D0CC3D84853BC0 CRC64;

Query Match 62.9%; Score 39; DB 3; Length 945;
Best Local Similarity 77.8%; Pred. No. 89;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 GSISANYVQ 13
||:|||||

DB 355 GSIVSNVYQ 363

RESULT 7

QYH4Y2 PRELIMINARY; PRT; 121 AA.

AC QYH4Y2
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE D180113.1 (Hemopoietic cell kinase) (Fragment).

GN HCK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;

RP SEQUENCE FROM N.A.

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.

DR EMBL: AL353092; CAC08002.1; -
DR HSSP: P08631; 1AD5
DR InterPro: IPR001018; Neu_cyt_fact_2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR00499; P67PHOX.
DR PRINTS: PR00452; SH3DOMAIN.
DR PROSITE: PS50002; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR KINASE: SH3 domain.
FT NON TER 121
SQ SEQUENCE 121 AA; 13363 MW; AA20F3132FBA1D9C CRC64;

Query Match 61.3%; Score 38; DB 4; Length 121;
Best Local Similarity: 66.7%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSISANYV 12
||:|||||

DB 101 TRKEGYIPSNV 112

RESULT 8

QYH4Y2 PRELIMINARY; PRT; 371 AA.

AC QYH4Y2
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG30004-PA.
GN CG30004.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;

RP SEQUENCE FROM N.A.

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AEO03832; AAM68791.1; -
DR FlyBase: FBgn0050004; CG30004.
SQ SEQUENCE 371 AA; 39670 MW; 6055C345D2EFA270 CRC64;

Query Match 61.3%; Score 38; DB 5; Length 371;
Best Local Similarity 58.3%; Pred. No. 49;

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).

[2] SEQUENCE FROM N.A.

RP Celinker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Buesam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorese V., Doup L.E., Doyle C., Dreenek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Paclet J., Paragaa V., Park S., Patel S., Pfeiffer B.,
RA Phoumenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svitek R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3] SEQUENCE FROM N.A.

RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[4] SEQUENCE FROM N.A.

RP Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;

[5] SEQUENCE FROM N.A.

RP Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO03832; AAM68791.1; -
DR FlyBase: FBgn0050004; CG30004.
SQ SEQUENCE 371 AA; 39670 MW; 6055C345D2EFA270 CRC64;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
||| : |||
217 TRSSNTSSSNYI 228

Db

RESULT 9

Q8AC27 PRELIMINARY; PRT; 498 AA.

AC 08AC27; 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)

DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Gag polyprotein precursor (Fragment).

GN GAG.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_Taxid=11676;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=98UG57143;

RX MEDLINE=20564795; PubMed=11112486;

RA Salminen M.O., Ehrenberg P.K., Mascola J.R., Dayhoff D.E., Merling R., Blake B., Louder M., Hegerich S., Polonis V.R., Bix D.L., Robb M.L., McCutchan F.E., Michael N.L.;

RT "Construction and biological characterization of infectious molecular clones of HIV-1 subtypes B and E (CRF01_AE) generated by the polymerase chain reaction.";

RT Virology 278:103-110(2000).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=98UG57143;

RX MEDLINE=22375625; PubMed=12487816;

RA Harris M.E., Serwadda D., Sewankambo N., Wabwire F., Kim B., Kiigizi G., Kiwanuka N., Phillips J.B., Meehan M., Lutalo T., Lane J.R., Merling R., Gray R., Wawer M., Bix D.L., Robb M.L., McCutchan F.E.;

RT "Among 46 Near Full length HIV Type 1 Genome Sequences from Rakai District, Uganda, Subtype D and AD Recombinants Predominate.";

RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).

RX EMBL, AF484514; AAN73764.1; -.

RA Polyprotein.

RT NON TER

FT

SQ SEQUENCE 498 AA; 55308 MW; FFC0A7094646CAd7 CRC64;

Query Match 61.3%; Score 38; DB 15; Length 498;
Best Local Similarity 54.5%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNY 11
||| : |||
120 TRSSGVSQNY 130

Db

RESULT 10

O13420 PRELIMINARY; PRT; 519 AA.

AC 013420; 01-JAN-1998 (TReMBLrel. 05, Created)

DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Phenoloxidase (EC 1.10.3.2).

GN FOX1.

OS Trimetes sp. I-62.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Trimetes.

OX NCBI_Taxid=51095;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=CECT 20197;

RX MEDLINE=97355933; PubMed=9212414;

RA Mansur M., Suarez T., Fernandez-Larrea J.B., Brizuela M.A.,

RA Gonzalez A.E.;

RT "Identification of a laccase gene family in the new lignin-degrading basidiomycete CECT 20197.";

RT Appl. Environ. Microbiol. 63:2637-2646(1997).

DR EMBL, U65399; AAB63443.1; -.

DR HSP, P37064; IAOZ.

DR InterPro, IPR001117; Cu-oxidase.

DR Pfam, PF00394; Cu-oxidase; 3.

DR PROSITE, PS00079; MULTICOPPER_OXIDASE1, 1.

OX Oxidoreductase.

SQ SEQUENCE 519 AA; 55665 MW; CD226C498770DDE4 CRC64;

Query Match 61.3%; Score 38; DB 3; Length 519;
Best Local Similarity 72.7%; Pred. No. 72;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RSSGSIASNYV 12
||| : |||
428 RSAGSTAYNV 438

Db

RESULT 11

O9KKL8 PRELIMINARY; PRT; 556 AA.

AC O9KKL8; 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)

DE Hypothetical protein VCA1085.

GN VCA1085.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrio.

OX NCBI_Taxid=666;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=E1 Tor N16961 / Serotype O1;

RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L., Dodson R.U., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tectel H., Richardson D., Ermolaeva M.D., Yamathavan J., Baes S., Qin H., Dragol I., Sellers P., McDonald L., Utecher T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";

RT Nature 406:477-483(2000).

RL EMBL, AB004434; AAF96978.1; -.

DR TIGR; VCA1085; -.

DR InterPro; IPR005646; DUF342.

DR Pfam; PF03961; DUF342; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 556 AA; 60476 MW; 8565B105BEF2E2DF CRC64;

Query Match 61.3%; Score 38; DB 16; Length 556;
Best Local Similarity 58.3%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RSSGSIASNYV 13
||| : |||
343 KSGSIRANVYQ 354

Db

RESULT 12

O9XXR3 PRELIMINARY; PRT; 606 AA.

AC O9XXR3; 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Y51A2D.4 protein.

GN Y51A2D.4.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OK NCBI_TaxID=6239;
 RN
 RA SEQUENCE FROM N.A.
 RP McMurtry A.A.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN
 RP
 RA SEQUENCE FROM N.A.
 MEDLINE=99069613; PubMed=9851916;
 none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RT Science 282:2012-2018(1998).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL, A021497; CA16400.1; -.
 DR WormPep; Y51A2D.4; CE19201.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003653; Sugar transporter.
 DR InterPro; IPR005828; Sug_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRfam; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transporter; transmembrane; Transport.
 KW SEQUENCE 606 AA; 66658 MW; 5A217A21AF87B501 CRC64;
 SQ

Query Match 61.3%; Score 38; DB 5; Length 606;
 Best Local Similarity 63.6%; Pred. No. 85;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNY 11
 Db 442 TNSGSIANNF 452

RESULT 13
 Q9V5B0 PRELIMINARY; PRT; 762 AA.
 AC Q9V5B0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE CG1776 protein.
 GN CG1776.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7227;
 RN
 RA SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY;
 MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitek R., Tector C., Turner R., Venter R., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenhach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RT Science 287:2185-2195 (2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AE003832; AAF58906.1; -.
 DR HSSP; G63450; 1A06.
 DR PDBase; P8N0033441; CG1776.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; PKinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 762 AA; 83776 MW; E447DFC9AEE68D CRC64;
 SQ

Query Match 61.3%; Score 38; DB 5; Length 762;
 Best Local Similarity 58.3%; Pred. No. 11e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNY 12
 Db 608 TRSSNTSSSNTY 619

RESULT 14
 Q8MMT7 PRELIMINARY; PRT; 1041 AA.
 AC Q8MMT7;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Microtubule-associated protein CP224. 6/101.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OK NCBI_TaxID=44689;
 RN
 RA SEQUENCE FROM N.A.
 RP STRAIN=MX4;
 RC Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
 RL Submitted (MAY-2002) to the EMBL/genbank/DBJ databases.
 DR EMBL; AC110767; AM33696.1; -.
 SQ SEQUENCE 1041 AA; 117567 MW; 5204E9AC93E0DC3 CRC64;
 SQ

Query Match 61.3%; Score 38; DB 5; Length 1041;
 Best Local Similarity 63.6%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNY 11

DB 355 TATSGSLGSNY 365

RESULT 15

Q9H3D8 PRELIMINARY; PRT; 1153 AA.
 AC Q9H3D8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE MSTP063.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RA Zhao B., Xu Y.Y., Liu Y.Q., Wang X.Y., Lui B., Ye J., Song L.,
 RA Zhao Y., Cao H.Q., Zhao X.W., Zhang C.L., Zhang J., Liu L.S.,
 RA Ding J.F., Gao R.L., Wu Q.Y., Qiang B.Q., Yuan J.G., Liew C.C.,
 RA Zhao M.S., Hui R.T.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF19814; AAG47945.1;
 DR InterPro; IPR007087; Znf_C2H2.
 DR SMART; SM00355; Znf_C2H2_1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 SQ SEQUENCE 1153 AA; 130800 MW; 327FEA675028D459 CRC64;

Query Match 61.3%; Score 38; DB 4; Length 1153;
 Best Local Similarity 58.3%; Pred. No. 1.8e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RSSGSIASNYVQ 13
 DB 108 KNSGSIKDYVR 119

Search completed: January 12, 2004, 07:02:08
 Job time : 31.4062 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:09 ; Search time 31.4844 Seconds
(without alignments)
65.539 Million cell updates/sec

Title: US-09-829-495-64
Perfect score: 62
Sequence: 1 TRSSGSIASNYVQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	13	AA61296	Anti-TANGO 268 scFv
2	62	100.0	13	ABU11262	Human anti-Fc-epsi
3	62	100.0	98	ABP56510	Human anti-Fc-epsi
4	62	100.0	110	AAO18434	Anti-GD2 antibody
5	62	100.0	113	AAO25222	Anti-adipocyte mon
6	62	100.0	161	AAO76011	Human colon cancer
7	58	93.5	248	ABU19832	Human VEGF-2 relat
8	56	90.3	258	ABP45223	Human Blys binding
9	55	88.7	112	ABP56513	Human anti-Fc-epsi

10	53	85.5	253	23	ABP44898	Human Blys binding
11	53	85.5	253	23	ABP45114	Human Blys binding
12	53	85.5	253	23	ABP45156	Human Blys binding
13	53	85.5	253	23	ABP45364	Human Blys binding
14	52	83.9	257	23	ABP45520	Human Blys binding
15	51	82.3	69	23	ABP62854	Human polypeptide
16	50	80.6	13	18	AAW16590	Anti-RSV F glycopr
17	50	80.6	109	18	AAW16588	Human monoclonal a
18	50	80.6	109	18	AAW14785	Human anti-Fc-epsi
19	45	72.6	114	24	ABP56511	Human anti-Fc-epsi
20	45	72.6	114	24	ABP56512	Human anti-Fc-epsi
21	44	71.0	677	22	ABG07372	Novel human diago
22	44	71.0	3572	23	ABG35659	Human nucleic acid
23	42	67.7	517	22	ABB57957	Drosophila melanog
24	39	62.9	13	13	AAK23736	Ligand CDR(d) . SY
25	39	62.9	13	14	AAK50108	CDR (d) . Syntheti
26	39	62.9	13	21	AAK68004	Amino acid sequenc
27	39	62.9	13	23	ABG31465	Rat immunoglobulin
28	39	62.9	110	23	ABG31476	Light chain variab
29	39	62.9	111	21	AAK68007	A light chain vari
30	39	62.9	128	13	AAK23772	Recombinant light
31	39	62.9	128	13	AAK23778	Recombinant light
32	39	62.9	128	13	AAK23814	Recombinant light
33	39	62.9	130	13	AAK23740	Recombinant light
34	39	62.9	130	13	AAK23775	Recombinant light
35	39	62.9	130	13	AAK23779	Recombinant light
36	39	62.9	130	13	AAK23817	Recombinant light
37	39	62.9	130	13	AAK23821	Recombinant light
38	39	62.9	130	13	AAK23823	Recombinant light
39	39	62.9	130	14	AAK41722	Aglycosylated anti
40	39	62.9	132	13	AAK23776	Recombinant light
41	39	62.9	132	13	AAK23822	Recombinant light
42	39	62.9	132	13	AAK23826	Recombinant light
43	39	62.9	134	13	AAK23769	Recombinant light
44	39	62.9	134	13	AAK23771	Recombinant light
45	39	62.9	134	13	AAK23777	Recombinant light

ALIGNMENTS

RESULT 1	AA61296	standard; Peptide; 13 AA.
ID	AA61296	
AC	AA61296;	
DT	04-APR-2001	(first entry)
DE	Anti-TANGO 268 scFv CDR, SEQ ID NO: 64.	
KW	Human; antibody; scFv; CDR; complementarity determining region;	
KW	TANGO 268; cardiant; cerebroprotective; cyostatic; antioedemant;	
KW	thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;	
KW	platelet membrane glycoprotein receptor; bleeding disorder;	
KW	blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;	
KW	ischaemia; cardiovascular disease; immunological disease; liver disorder;	
KW	cancer.	
OS	Homo sapiens.	
PN	WO200100810-A1.	
PD	04-JAN-2001.	
PF	30-JUN-2000; 2000WO-US18152.	
PR	30-JUN-1999; 99US-0345468.	
PR	06-DEC-1999; 99US-0454824.	
PR	14-FEB-2000; 2000US-0503387.	
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		

PI Busfield SJ, Villelail J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI Qian DM, Kingsbury G;
XX
XX
DR MPI; 2001-080877/09.
XX
XX PT New genes encoding human platelet-expressed collagen receptor,
PT glycoprotein VI, and its modulators, useful for preventing, treating
PT and diagnosing hemorrhagic disorders, thrombotic diseases and
PT immunological disorders -
XX
XX PS Claim 32; Page 102; 227pp; English.
XX
XX CC The present sequence is given in a specification relating to an isolated
CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
CC and polypeptides and their modulators, e.g. antisense nucleic acids,
CC ribozymes and antibodies, are useful for preventing, treating and
CC diagnosing disorders associated with aberrant expression or activity of
CC GPVI. These disorders include bleeding disorders
CC (e.g. thrombocytopenia), blood vessel injury, thrombotic disorders
CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC infarction), immunological diseases (e.g. platelet disorder) and
CC embryonic liver disorders. Preferably they are used to prevent acute
CC cardiac ischaemia following angioplasty and metastatic cancers,
CC especially of the colon and liver.
XX
XX SQ Sequence 13 AA;
XX
XX Query Match 100.0%; Score 62; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TRSSGSIASNYVQ 13
Db 1 TRSSGSIASNYVQ 13
XX
XX
XX RESULT 2
ID ABU11262 standard; Peptide; 13 AA.
XX
XX AC ABU11262;
XX
XX DT 06-FEB-2003 (first entry)
XX
XX DE Human TANGO 268 VLCDRI Peptide #3.
XX
XX KW Human; mouse; variable heavy; VH; antigen; cancer;
KW complementarity determining region; TANGO 268; glycoprotein VI; GPVI;
KW TANGO 268; extracellular matrix; collagen; platelet release;
KW proliferation; migration; embryogenesis; inflammation; thrombosis;
KW degradation; thrombocytopenia; antibody; thrombotic disorder;
KW cerebral vascular disease; stroke; ischaemia; venous thromboembolism;
KW leg swelling; pain; ulceration; pulmonary embolism; coronary disease;
KW cardiovascular disease; angina pectoris; myocardial infarction;
KW coronary restenosis; atherosclerosis; immunological disorder;
KW developmental disorder; embryonic disorder; liver disorder;
KW cerebral vascular disease; venous thromboembolism disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200280968-A1.
XX
XX PD 17-OCT-2002.
XX
XX PF 09-APR-2002; 2002WO-US11122.
XX
XX PR 09-APR-2001; 2001US-0829495.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX

PI Busfield SJ, Villelail J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI Qian DM, Kingsbury G;
XX
XX
DR MPI; 2003-058477/05.
XX
XX PT Novel substantially purified antibody immunospecifically binding to
PT TANGO 268 antigen, useful for treating bleeding disorders such as
PT thrombocytopenia, stroke, ischemia, pulmonary embolism, atherosclerosis
PT -
XX
XX PS Claim 10; Page 111; 236pp; English.
XX
XX CC This invention relates to a novel purified antibody comprising a
CC variable heavy (VH) complementarity determining region (CDR1, VH CDR2
CC or VH CDR3; or variable light (VL) CDR1, VL CDR2 or VL CDR3, and
CC immunospecifically binding to a TANGO 268 (also referred as glycoprotein
CC VI (GPVI)) antigen. The antibodies of the invention act to decrease or
CC block TANGO 268 binding to extracellular matrix components, or as a
CC collagen or platelet release and aggregation blocker. The antibodies of
CC the invention are useful for modulating proliferation, migration,
CC morphology, differentiation and/or function of megakaryocytes and
CC platelets, including during development e.g. embryogenesis, modulating
CC leukocyte-platelet and platelet-endothelium interactions in
CC inflammation and/or thrombosis, and modulating platelet aggregation and
CC degradation. They are also useful for modulating disorders associated
CC with abnormal or aberrant megakaryocyte and/or platelet proliferation,
CC migration, morphology, differentiation and/or function, e.g. bleeding
CC disorders such as thrombocytopenia. Other diseases which may be
CC modulated by these antibodies are thrombotic disorders, cerebral
CC vascular diseases (e.g. stroke and ischaemia) venous thromboembolism
CC diseases (e.g. diseases involving leg swelling, pain and ulceration,
CC pulmonary embolism, etc); coronary diseases (e.g. cardiovascular
CC diseases including angina pectoris, myocardial infarction, coronary
CC restenosis, atherosclerosis, etc); immunological disorders,
CC developmental disorders, embryonic disorders, liver disorders, cerebral
CC vascular diseases, venous thromboembolism disease, coronary diseases,
CC and metastatic cancers. The antibodies of the invention only causes a
CC transient decrease in platelet counts, platelet aggregation, and/or
CC platelet activation and so have some advantages over prior art
CC methods. The present sequence represents a peptide sequence used to
CC generate the antibodies of the invention.
XX
XX SQ Sequence 13 AA;
XX
XX Query Match 100.0%; Score 62; DB 24; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TRSSGSIASNYVQ 13
Db 1 TRSSGSIASNYVQ 13
XX
XX
XX RESULT 3
ID ABP56510 standard; Protein; 98 AA.
XX
XX AC ABP56510;
XX
XX DT 20-MAR-2003 (first entry)
XX
XX DE Human anti-Fc-epsilon-R1 alpha autoantibody light chain VI-22.
XX
XX KW Autoantibody; Fc-epsilon-R1 receptor alpha-chain; immunosuppressive;
KW allergic disease; urticaria; late phase allergic reaction; malignancy;
KW intrinsic asthma; drug intolerance; food intolerance; immunoglobulin E;
KW conditional autoimmunity; IGE mediated disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200282085-A2.
XX

PD 17-OCT-2002.
XX
XX 03-APR-2002; 2002WO-EP03660.
XX
XX 04-APR-2001; 2001US-281024P.
XX
XX (ZLBB-) ZLB BIOPLASMA AG.
XX
XX Miescher S;
XX
XX WPI; 2003-103348/09.
DR
XX Identifying and obtaining inhibitor of a pathological process for
PT treating e.g. autoimmunity comprises determining if a compound is
PT capable of modulating the binding of the Fc-epsilon-R1 receptor and an
PT autoantibody against its alpha-chain -
XX
XX Claim 20; Page 22; 29pp; English.
XX
XX The present invention describes a method for identifying and obtaining
CC an inhibitor of a pathological process. The method comprises determining
CC if a compound is capable of modulating the binding of the Fc-epsilon-R1
CC receptor alpha-chain and an autoantibody against its alpha-chain. Also
CC described: (1) use of the autoantibody against the Fc-epsilon-R1 receptor
CC alpha-chain for identifying and obtaining an inhibitor of a pathological
CC process; (2) use of the identified inhibitor for inhibiting activity of
CC the autoantibody against the Fc-epsilon-R1 receptor alpha-chain; and
CC (3) a compound identified by the method, which binds but does not
CC activate the receptor; and (4) a polypeptide capable of specific binding
CC to the Fc-epsilon-R1 receptor alpha-chain. The method is useful for
CC obtaining an inhibitor of a pathological process e.g. imbalance between
CC cell-bound and free IGE e.g. allergic disease (urticaria, late phase
CC allergic reactions, intrinsic asthma, drug intolerance and food
CC intolerance), IGE mediated disease or malignancy. The compound is useful
CC for treating a pathological process, particularly conditional
CC autoimmunity. The present sequence represents a human recombinant
CC anti-Fc-epsilon-R1 alpha autoantibody light chain protein sequence from
CC the present invention.
XX
XX
SQ Sequence 98 AA;
Query Match 100.0%; Score 62; DB 24; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYVQ 13
DB 23 TRSSGSIASNYVQ 35
RESULT 4
AA018434
ID AA018434 standard; protein; 110 AA.
XX
XX AA018434;
XX
XX 11-OCT-2002 (first entry)
XX
XX Anti-GD2 antibody light chain #11.
XX
XX Anti-GD2 antibody; neuroectodermal tumour; cancer; neuroblastoma;
XX melanoma; vaccine; gene therapy; cytostatic; disialoganglioside GD2.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX DE10059930-A1.
XX
XX 29-MAY-2002.
XX
XX 23-NOV-2000; 2000DE-1059930.
XX
XX 23-NOV-2000; 2000DE-1059930.
PR

XX
XX (FISC/) FISCHER P.
XX PA (UTTE/) UTTENEUTHER-FISCHER M.
XX
XX Uttenreuther-Fischer M, Krueger J;
XX
XX WPI; 2002-51006/55.
XX
XX Composition for treating tumors that produce disialoganglioside GD2,
PT comprises human antibody fragment able to induce anti-idiotypic
PT antibodies -
XX
XX Claim 3; Fig 7A; 14pp; German.
XX
XX The present invention relates to a composition for treating tumors
CC positive for disialoganglioside GD2 which is based on human antibody
CC fragments that activate the immune system against GD2, specifically by
CC inducing anti-GD2 antibodies. The composition can be used in the
CC treatment of neuroblastoma and melanoma. The present sequence is an
CC anti-GD2 antibody light chain.
XX
XX
SQ Sequence 110 AA;
Query Match 100.0%; Score 62; DB 23; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.00074;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYVQ 13
DB 21 TRSSGSIASNYVQ 33
RESULT 5
AA02522
ID AA02522 standard; protein; 113 AA.
XX
XX AA02522;
XX
XX 29-AUG-2001 (first entry)
XX
XX
DE Anti-adipocyte monoclonal antibody light chain, FAT 13.
XX
XX Antibody; adipocyte; heavy chain; light chain; obesity; fat;
XX heart disease; complementarity determining region; CDR.
XX
XX Homo sapiens.
XX
XX WO200127279-A1.
XX
XX 19-APR-2001.
XX
XX 11-OCT-2000; 2000WO-GB03900.
XX
XX 12-OCT-1999; 99US-0158812.
XX
XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Edwards BM, Main SH, Vaughan TJ;
XX
XX WPI; 2001-282031/29.
XX
XX N-PSDB; AAS03422.
XX
XX Panel of specific binding members of antibody molecules which bind to
PT whole adipocytes is used in the treatment of obesity and obesity
PT related diseases -
XX
XX
XX Claim 1; Page 106; 182pp; English.
XX
XX AA02501-AA02635, and AA02641-AA02748 represent the amino acid
CC sequences of anti-adipocyte monoclonal antibody heavy chain, light
CC chain, and heavy chain complementarity determining regions (CDR) of the
CC invention. The antibodies can be used in the treatment of obesity and
CC obesity related diseases. The antibodies can be used to deliver drugs or

CC pro-drugs directly to the fat mass of an obese patient or the antibody
CC can be used as a therapeutic itself. Antibodies binding specifically to
CC adipocytes can be used to activate the immune system to destroy the cells
CC by complement mediated lysis. The antibodies may be labeled with a
CC detectable label such as radiolabel, fluorescent or chemical group and
CC used in methods of diagnosis in human subjects e.g. to determine the
CC presence of adipocyte antigen on the surface of an adipocyte to detect or
CC determine the presence or level of adipocytes in a cell or tissue sample.
CC The antibodies can be used as an alternative means of treatment for obese
CC patients other than undergoing surgery to remove excess fat. Antibodies
CC for different types of fat deposits can also be produced e.g. intra-
CC abdominal fat associated with heart disease.

CC Sequence 113 AA;

Query Match 100.0%; Score 62; DB 22; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
| | | | | | | | | | | | | | |
Db 25 TRSSGSIASNYVQ 37

RESULT 6

ID AAG76011 standard; Protein; 161 AA.

AC AAG76011;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:6775.

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

KM colorectal carcinoma; chromosome 22.

XX Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI, 2001-235357/24.

DR N-PSDB; AAH35416.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 8234-8235; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis

CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

CC Sequence 161 AA;

Query Match 100.0%; Score 62; DB 22; Length 161;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
| | | | | | | | | | | | | | |
Db 44 TRSSGSIASNYVQ 56

RESULT 7

ID ABJ19832 standard; Protein; 248 AA.

AC ABJ19832;

DT 10-APR-2003 (first entry)

DE Human VEGF-2 related protein SEQ ID NO 78.

KW Cytostatic; cardiac; cardiovascular; antiinflammatory; antirheumatic;

KW antidiabetic; antidiabetic; ophthalmological; antiallergic; antibody;

KW immunosuppressive; dermatological; antiproliferative; vulnary; cancer;

KW CDR region; VH domain; VL domain; immunospecific; VEGF-2; Cancer;

KW proliferative disorder; cardiovascular disorder; arrhythmia;

KW cerebrovascular disorder; cerebral anoxia; inflammatory disease;

KW infectious disease; autoimmune disease; rheumatoid arthritis;

KW systemic lupus erythematosus; allergy; diabetic retinopathy; psoriasis;

KW angiogenesis; wound healing; vascular tissue repair; human.

XX Homo sapiens.

PN WO200283704-A1.

PD 24-OCT-2002.

PF 12-APR-2002; 2002WO-US11474.

PR 13-APR-2001; 2001US-283385P.

PR 24-JAN-2002; 2002US-350366P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Albert VR, Ruben SM, Wager RE;

DR WPI, 2003-092991/08.

PT New isolated polynucleotide encoding an antibody which inhibits a
PT VEGF-2 polypeptide, useful for diagnosing, treating or preventing
PT diseases associated with aberrant VEGF-2 expression or function, e.g.
PT cancer or inflammation -

XX Disclosure; Page 389-390; 425pp; English.

CC The invention relates to an isolated polynucleotide encoding a first
CC antibody at least 95-100% identical to a second antibody comprising an
CC amino acid sequence selected from at least one, two or three CDR
CC region(s) of a VH or VL domain where the first antibody
CC immunospecifically inhibits a VEGF-2 polypeptide. The isolated
CC polynucleotide is useful in diagnosing, treating, preventing, prognosising,
CC ameliorating or monitoring diseases associated with aberrant VEGF-2 or
CC VEGF-2 receptor expression or lack of VEGF-2 or VEGF-2 receptor function,
CC such as cancer and other proliferative disorders, cardiovascular
CC disorders (arrhythmias), cerebrovascular disorders (e.g. cerebral
CC anoxia), inflammatory diseases, infectious diseases, autoimmune diseases

(e.g. rheumatoid arthritis, Systemic Lupus Erythematosus, allergies),
diabetic retinopathy or psoriasis. The polynucleotide, polypeptide and
antibodies may also be used to stimulate angiogenesis, wound healing, and
promoting vascular tissue repair. The polynucleotide and polypeptide may
also be used for in vitro purposes related to scientific research.
CC synthesis of DNA and manufacture of DNA vectors, and for the production
of diagnostics and therapeutics to treat human diseases. This sequence
CC represents a human VEGF-2 related protein of the invention.

Sequence 248 AA;

Query Match 93.5%; Score 58; DB 24; Length 248;
Best Local Similarity 92.3%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TRSSGSIASNYVQ 13
|||
Db 160 TRSSGSIASNYVQ 172

RESULT 8
ABP45223

ID ABP45223 standard; Protein; 258 AA.

XX ABP45223;

19-AUG-2002 (first entry)

Human Blys binding scFv SEQ ID 1234.

Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
tumour necrosis factor; B cell proliferation; B cell differentiation;
immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
common variable immunodeficiency; acquired immunodeficiency syndrome.

Homo sapiens.

WO200202641-A1.

10-JAN-2002.

15-JUN-2001; 2001WO-US19110.

16-JUN-2000; 2000US-212210P.

17-OCT-2000; 2000US-240816P.

16-MAR-2001; 2001US-276248P.

21-MAR-2001; 2001US-277379P.

25-MAY-2001; 2001US-293499P.

(HUMA-) HUMAN GENOME SCI INC.

(CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

WPI; 2002-114799/15.

Antibodies against B lymphocyte Stimulating polypeptides, useful for
the diagnosis and treatment of cancers and immune disorders -

Claim 1; Page 1874-1875; 3148pp; English.

This invention describes novel antibodies that immunospecifically bind to
B lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
tumour necrosis factor (TNF) super family and induces B cell
proliferation and differentiation. The antibodies of the invention have
cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
antiinflammatory and antiAIDS activity and can be used in vaccines to
inhibit the expression and activity of Blys. The antibodies bind to Blys
and so may be used to detect and quantitate the presence of Blys in
biological samples and may be used in this way to diagnose disease
associated with aberrant expression of Blys. They may also be

administered to treat diseases associated with aberrant Blys expression
and activity such as cancer, immune, and autoimmune disorders and
diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
immunodeficiency (e.g. common variable immunodeficiency (CVID) and
acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
the antibodies and fragments of the antibodies described in the method
of the invention.

Sequence 258 AA;

Query Match 90.3%; Score 56; DB 23; Length 258;
Best Local Similarity 92.3%; Pred. No. 0.027;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TRSSGSIASNYVQ 13
|||
Db 170 TRSSGSIDSNYVQ 182

RESULT 9
ABP56513

ID ABP56513 standard; Protein; 112 AA.

XX ABP56513;

20-MAR-2003 (first entry)

Human anti-Fc-epsilon-R1 alpha autoantibody light chain UG-alpha-8.

Autoantibody; Fc-epsilon-R1 receptor alpha-chain; immunosuppressive;
allergic disease; urticaria; late phase allergic reaction; malignancy;
intrinsic asthma; drug intolerance; food intolerance; immunoglobulin E;
conditional autoimmunity; IGE mediated disease.

Homo sapiens.

Synthetic.

WO200282085-A2.

17-OCT-2002.

03-APR-2002; 2002WO-EP03660.

04-APR-2001; 2001US-281024P.

(ZLBB-) ZLB BIOPHARMA AG.

Miescher S;

WPI; 2003-103348/09.

Identifying and obtaining inhibitor of a pathological process for
treating e.g. autoimmunity comprises determining if a compound is
capable of modulating the binding of the Fc-epsilon-R1 receptor and an
autoantibody against its alpha-chain -

Claim 20; Page 22; 29pp; English.

The present invention describes a method for identifying and obtaining
an inhibitor of a pathological process. The method comprises determining
if a compound is capable of modulating the binding of the Fc-epsilon-R1
receptor alpha-chain and an autoantibody against its alpha-chain. Also
described: (1) use of the autoantibody against the Fc-epsilon-R1 receptor
alpha-chain for identifying and obtaining an inhibitor of a pathological
process; (2) use of the identified inhibitor for inhibiting activity of
the autoantibody against the Fc-epsilon-R1 receptor alpha-chain; and
(3) a compound identified by the method, which binds but does not
activate the receptor; and (4) a polypeptide capable of specific binding
to the Fc-epsilon-R1 receptor alpha-chain. The method is useful for
obtaining an inhibitor of a pathological process e.g. imbalance between
cell-bound and free IGE e.g. allergic disease (urticaria, late phase
allergic reactions, intrinsic asthma, drug intolerance and food
intolerance), IGE mediated disease or malignancy. The compound is useful

CC for treating a pathological process, particularly conditional
CC autoimmunity. The present sequence represents a human recombinant
CC anti-Fc-epsilon-R1 alpha autoantibody light chain protein sequence from
CC the present invention.

XX Sequence 112 AA;

Query Match 88.7%; Score 55; DB 24; Length 112;
Best Local Similarity 92.3%; Pred. No. 0.016;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
Db 21 TRSSDSIASNYVQ 33

RESULT 10

ABP44898
ID ABP44898 standard; Protein; 253 AA.

XX AC ABP44898;

XX DT 19-AUG-2002 (first entry)

XX DE Human BlyS binding scFv SEQ ID 909.

XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytosolic;
XX tumour necrosis factor; B cell proliferation; B cell differentiation;
XX immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
XX antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
XX systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
XX common variable immunodeficiency; acquired immunodeficiency syndrome.

OS Homo sapiens.

XX PN WO200202641-A1.

XX PD 10-JAN-2002.

XX PF 15-JUN-2001; 2001WO-US19110.

XX PR 16-JUN-2000; 2000US-212210P.

XX PR 17-OCT-2000; 2000US-240816P.

XX PR 16-MAR-2001; 2001US-276248P.

XX PR 21-MAR-2001; 2001US-277379P.

XX PR 25-MAY-2001; 2001US-293499P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX DR WPI; 2002-114799/15.

XX PT Antibodies against B lymphocyte stimulating polypeptides, useful for
XX the diagnosis and treatment of cancers and immune disorders -

XX PS Claim 1; Page 1486-1487; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to
XX B lymphocyte stimulator (BlyS) polypeptides. BlyS is a member of the
XX tumour necrosis factor (TNF) super family and induces B cell
XX proliferation and differentiation. The antibodies of the invention have
XX cytosolic, immunosuppressive, immunostimulant, immunomodulatory,
XX antirheumatic and antiAIDS activity and can be used in vaccines to
XX inhibit the expression and activity of BlyS. The antibodies bind to BlyS
XX and so may be used to detect and quantitate the presence of BlyS in
XX biological samples and may be used in this way to diagnose disease
XX associated with aberrant expression of BlyS. They may also be
XX administered to treat diseases associated with aberrant BlyS expression
XX and activity such as cancer, immune, and autoimmune disorders and
XX diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
XX immunodeficiency (e.g. common variable immunodeficiency (CVID) and

CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method
CC of the invention.

XX Sequence 253 AA;

Query Match 85.5%; Score 53; DB 23; Length 253;
Best Local Similarity 84.6%; Pred. No. 0.095;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
Db 164 TRSSGSIASNYVQ 176

RESULT 11

ABP45114
ID ABP45114 standard; Protein; 253 AA.

XX AC ABP45114;

XX DT 19-AUG-2002 (first entry)

XX DE Human BlyS binding scFv SEQ ID 1125.

XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytosolic;
XX tumour necrosis factor; B cell proliferation; B cell differentiation;
XX immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
XX antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
XX systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
XX common variable immunodeficiency; acquired immunodeficiency syndrome.

OS Homo sapiens.

XX PN WO200202641-A1.

XX PD 10-JAN-2002.

XX PF 15-JUN-2001; 2001WO-US19110.

XX PR 16-JUN-2000; 2000US-212210P.

XX PR 17-OCT-2000; 2000US-240816P.

XX PR 16-MAR-2001; 2001US-276248P.

XX PR 21-MAR-2001; 2001US-277379P.

XX PR 25-MAY-2001; 2001US-293499P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX DR WPI; 2002-114799/15.

XX PT Antibodies against B lymphocyte stimulating polypeptides, useful for
XX the diagnosis and treatment of cancers and immune disorders -

XX PS Claim 1; Page 1745-1746; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to
XX B lymphocyte stimulator (BlyS) polypeptides. BlyS is a member of the
XX tumour necrosis factor (TNF) super family and induces B cell
XX proliferation and differentiation. The antibodies of the invention have
XX cytosolic, immunosuppressive, immunostimulant, immunomodulatory,
XX antirheumatic and antiAIDS activity and can be used in vaccines to
XX inhibit the expression and activity of BlyS. The antibodies bind to BlyS
XX and so may be used to detect and quantitate the presence of BlyS in
XX biological samples and may be used in this way to diagnose disease
XX associated with aberrant expression of BlyS. They may also be
XX administered to treat diseases associated with aberrant BlyS expression
XX and activity such as cancer, immune, and autoimmune disorders and
XX diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
XX immunodeficiency (e.g. common variable immunodeficiency (CVID) and
XX acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent

CC the antibodies and fragments of the antibodies described in the method
CC of the invention.

SQ Sequence 253 AA;

Query Match 85.5%; Score 53; DB 23; Length 253;
Best Local Similarity 84.6%; Pred. No. 0.095;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 TRSGSISASNYQ 13
|||||:|||||
Db 164 TRSGNISASKYQ 176

RESULT 12

ABP45156
ID ABP45156 standard; Protein; 253 AA.

AC ABP45156;

DT 19-AUG-2002 (first entry)

DE Human Blys binding scFv SEQ ID 1167.

KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.

OS Homo sapiens.

PN WO200202641-A1.

PD 10-JAN-2002.

PF 15-JUN-2001; 2001WO-US19110.

PR 16-JUN-2000; 2000US-212210P.

PR 17-OCT-2000; 2000US-240816P.

PR 16-MAR-2001; 2001US-276248P.

PR 21-MAR-2001; 2001US-277379P.

PR 25-MAY-2001; 2001US-293499P.

(HUMA-) HUMAN GENOME SCI INC.
(CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

DR WPI; 2002-114799/15.

PT Antibodies against B lymphocyte Stimulating polypeptides, useful for
the diagnosis and treatment of cancers and immune disorders -

Claim 1; Page 1795-1796; 3148pp; English.

CC This invention describes novel antibodies that immunospecifically bind to
CC B lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of Blys. The antibodies bind to Blys
CC and so may be used to detect and quantitate the presence of Blys in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of Blys. They may also be
CC administered to treat diseases associated with aberrant Blys expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method

CC of the invention.

SQ Sequence 253 AA;

Query Match 85.5%; Score 53; DB 23; Length 253;
Best Local Similarity 84.6%; Pred. No. 0.095;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 TRSGSISASNYQ 13
|||||:|||||
Db 164 TRSGNISASKYQ 176

RESULT 13

ABP45364
ID ABP45364 standard; Protein; 253 AA.

AC ABP45364;

DT 19-AUG-2002 (first entry)

DE Human Blys binding scFv SEQ ID 1375.

KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.

OS Homo sapiens.

PN WO200202641-A1.

PD 10-JAN-2002.

PF 15-JUN-2001; 2001WO-US19110.

PR 16-JUN-2000; 2000US-212210P.

PR 17-OCT-2000; 2000US-240816P.

PR 16-MAR-2001; 2001US-276248P.

PR 21-MAR-2001; 2001US-277379P.

PR 25-MAY-2001; 2001US-293499P.

(HUMA-) HUMAN GENOME SCI INC.
(CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

DR WPI; 2002-114799/15.

PT Antibodies against B lymphocyte Stimulating polypeptides, useful for
the diagnosis and treatment of cancers and immune disorders -

Claim 1; Page 2043-2044; 3148pp; English.

CC This invention describes novel antibodies that immunospecifically bind to
CC B lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of Blys. The antibodies bind to Blys
CC and so may be used to detect and quantitate the presence of Blys in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of Blys. They may also be
CC administered to treat diseases associated with aberrant Blys expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method

XX SQ Sequence 253 AA;
Query Match 85.5%; Score 53; DB 23; Length 253;
Best Local Similarity 84.6%; Pred. No. 0.095; 1; Mismatches 1; Indels 0; Gaps 0;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYVQ 13
| | | | | | | | | | | | | | |
Db 164 TRSSGSIASNYVQ 176

RESULT 14
ABP45520
ID ABP45520 standard; Protein; 257 AA.
XX AC ABP45520;
XX DT 19-AUG-2002 (first entry)
XX DE Human BlyS binding scfv SEQ ID 1531.
XX KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
XX KW tumour necrosis factor; B cell proliferation; B cell differentiation;
XX KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
XX KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
XX KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
XX KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX OS Homo sapiens.
XX PN WO200202641-A1.
XX PD 10-JAN-2002.
XX PF 15-JUN-2001; 2001WO-US19110.
XX PR 16-JUN-2000; 2000US-212210P.
XX PR 17-OCT-2000; 2000US-240816P.
XX PR 16-MAR-2001; 2001US-276248P.
XX PR 21-MAR-2001; 2001US-277379P.
XX PR 25-MAY-2001; 2001US-293499P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX DR WPI; 2002-114799/15.
XX PT Antibodies against B lymphocyte stimulating polypeptides, useful for
XX PT the diagnosis and treatment of cancers and immune disorders -
XX PS Claim 1; Page 2230-2231; 3148pp; English.
XX XX This invention describes novel antibodies that immunospecifically bind to
XX CC B lymphocyte stimulator (BlyS) polypeptides. BlyS is a member of the
XX CC tumour necrosis factor (TNF) super family and induces B cell
XX CC proliferation and differentiation. The antibodies of the invention have
XX CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
XX CC antirheumatic and antiAIDS activity and can be used in vaccines to
XX CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
XX CC and so may be used to detect and quantitate the presence of BlyS in
XX CC biological samples and may be used in this way to diagnose disease
XX CC associated with aberrant expression of BlyS. They may also be
XX CC administered to treat diseases associated with aberrant BlyS expression
XX CC and activity such as cancer, immune, and autoimmune disorders and
XX CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
XX CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
XX CC acquired immunodeficiency syndrome (AIDS)). ABP4590-ABP4728 represent
XX CC of the invention.

XX SQ Sequence 257 AA;
Query Match 83.9%; Score 52; DB 23; Length 257;
Best Local Similarity 84.6%; Pred. No. 0.15; 2; Mismatches 2; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYVQ 13
| | | | | | | | | | | | | | |
Db 167 TRSSGSIASNYVQ 179

RESULT 15
ABP62854
ID ABP62854 standard; Protein; 69 AA.
XX AC ABP62854;
XX DT 14-OCT-2002 (first entry)
XX DE Human polypeptide SEQ ID NO 291.
XX KW Human; vulnery; dermatological; neuroprotective; nootropic; cancer;
XX KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
XX KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
XX KW burn; central nervous system disorder; Alzheimer's disease;
XX KW Parkinson's disease; Huntington's disease; immune disorder;
XX KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX OS Homo sapiens.
XX PN WO200218424-A2.
XX PD 07-MAR-2002.
XX PF 31-AUG-2001; 2001WO-US27093.
XX PR 01-SEP-2000; 2000US-0654935.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
XX PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX DR WPI; 2002-583321/62.
XX DR N-PSDB; ABQ93333.
XX PT New polynucleotide and polypeptides, useful for treatment and diagnosis
XX PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
XX PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
XX PT sclerosis, diabetes and allergies -
XX PS Claim 20; SEQ ID NO 291; 284pp + Sequence Listing; English.
XX XX The invention relates to an isolated polynucleotide (I) comprising one of
XX CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
XX CC administering to a mammalian subject a composition comprising the protein
XX CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
XX CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
XX CC (I) is useful for gene therapy of diseases and (II) can be used for
XX CC therapeutic treatment. Diseases that may be treated include wound healing
XX CC and tissue repair, burns, central nervous system disorders (e.g.
XX CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
XX CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
XX CC sclerosis, diabetes and allergies.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pat_sequences.

XX SQ Sequence 69 AA;
Query Match 82.3%; Score 51; DB 23; Length 69;
Best Local Similarity 84.6%; Pred. No. 0.05; 1; Mismatches 1; Indels 0; Gaps 0;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRSSGSIASNYVQ 13
|||:||||
Db 42 TRSSGSIASNYVQ 54

Search completed: January 12, 2004, 06:58:56
Job time : 33.4844 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 07:02:25 ; Search time 26 Seconds
(without alignments)
100.755 Million cell updates/sec

Title: US-09-829-495-64

Perfect score: 62
Sequence: 1 TRSSGSIASNYVQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	13	9	US-09-832-312-64
2	62	100.0	13	12	US-09-829-495-64
3	62	100.0	98	12	US-10-308-817-117
4	62	100.0	110	15	US-10-269-805-50
5	62	100.0	110	15	US-10-269-805-52
6	62	100.0	161	15	US-10-106-658-6785
7	59	95.2	110	15	US-10-269-805-42
8	58	93.5	258	11	US-10-120-414-78
9	56	90.3	13	11	US-09-880-748-1234
10	55	88.7	13	11	US-09-972-656-2
11	55	88.7	216	11	US-09-972-656-108
12	54	87.1	13	11	US-09-972-656-3
13	54	87.1	216	11	US-09-972-656-90
14	53	85.5	253	11	US-09-880-748-909
15	53	85.5	253	11	US-09-880-748-1125

16	53	85.5	253	11	US-09-880-748-1167	Sequence 1167, Ap
17	53	85.5	253	11	US-09-880-748-1375	Sequence 1375, Ap
18	52	83.9	257	11	US-09-880-748-1531	Sequence 1531, Ap
19	51	82.3	110	15	US-10-269-805-58	Sequence 58, Appl
20	50	80.6	13	11	US-09-972-656-1	Sequence 1, Appl
21	50	80.6	103	11	US-09-972-656-130	Sequence 130, App
22	50	80.6	217	11	US-09-972-656-88	Sequence 88, Appl
23	49	79.0	243	12	US-10-322-673-55	Sequence 55, Appl
24	47	75.8	13	11	US-09-972-656-9	Sequence 9, Appl
25	47	75.8	218	11	US-09-972-656-102	Sequence 102, App
26	39	62.9	13	10	US-09-736-371B-8	Sequence 8, Appl
27	39	62.9	13	14	US-10-060-714-4	Sequence 4, Appl
28	39	62.9	110	14	US-10-060-714-15	Sequence 15, Appl
29	39	62.9	111	10	US-09-736-371B-15	Sequence 15, Appl
30	39	62.9	216	10	US-09-736-371B-19	Sequence 19, Appl
31	38	61.3	505	9	US-09-972-269-17	Sequence 17, Appl
32	38	61.3	505	10	US-09-972-269-17	Sequence 17, Appl
33	38	61.3	505	11	US-09-972-269-17	Sequence 17, Appl
34	38	61.3	526	12	US-10-394-322A-31	Sequence 31, Appl
35	38	61.3	606	12	US-10-369-493-6673	Sequence 6673, Ap
36	38	61.3	767	12	US-10-094-749-2369	Sequence 2369, Ap
37	36	58.1	109	10	US-09-144-886-91	Sequence 91, Appl
38	36	58.1	110	15	US-10-269-805-40	Sequence 40, Appl
39	36	58.1	352	12	US-10-369-493-5682	Sequence 5682, Ap
40	36	58.1	522	10	US-09-817-764-3	Sequence 3, Appl
41	36	58.1	522	12	US-10-253-904-49	Sequence 49, Appl
42	36	58.1	522	12	US-10-392-970-49	Sequence 49, Appl
43	35	56.5	58	9	US-09-879-957-131	Sequence 131, App
44	35	56.5	174	9	US-09-975-901-18	Sequence 18, Appl
45	35	56.5	181	9	US-09-975-901-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-832-312-64
; Sequence 64, Application US/09832312
; Patent No. US2001004982A1
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-832-312-64

Query Match
Best Local Similarity 100.0%; Score 62; DB 9; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|||||
1 TRSSGSIASNYVQ 13

Db 1 TRSSGSIASNYVQ 13

RESULT 2
US-09-829-495-64
; Sequence 64, Application US/09829495

```
/ Publication No. US20040001826A1
/ GENERAL INFORMATION:
/ APPLICANT: Busfield, S.J.
/ APPLICANT: Villeval J
/ APPLICANT: Jandrot-Perrus M
/ APPLICANT: Vainchenker W
/ APPLICANT: G111 DS
/ APPLICANT: Olan MD
/ TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
/ FILE REFERENCE: 7853-234
/ CURRENT APPLICATION NUMBER: US/09/829,495
/ CURRENT FILING DATE: 2001-04-09
/ PRIOR APPLICATION NUMBER: 09/610,118
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: 09/503,387
/ PRIOR FILING DATE: 2000-02-14
/ PRIOR APPLICATION NUMBER: 09/454,824
/ PRIOR FILING DATE: 1999-12-06
/ PRIOR APPLICATION NUMBER: 09/345,468
/ PRIOR FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 64
/ LENGTH: 13
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-829-495-64
```

```
Query Match          100.0%; Score 62; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TRSSGSIASNYVQ 13
        |||||
Db       1 TRSSGSIASNYVQ 13
```

```
RESULT 3
US-10-308-817-117
/ Sequence 117, Application US/10308817
/ Publication No. US20030219861A1
/ GENERAL INFORMATION:
/ APPLICANT: Roher, Russell
/ APPLICANT: Wu, Dayang
/ TITLE OF INVENTION: HYBRID ANTIBODIES
/ FILE REFERENCE: 1087-37
/ CURRENT APPLICATION NUMBER: US/10/308,817
/ CURRENT FILING DATE: 2002-12-03
/ NUMBER OF SEQ ID NOS: 195
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 117
/ LENGTH: 98
/ TYPE: PRT
/ ORGANISM: human
US-10-308-817-117
```

```
Query Match          100.0%; Score 62; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TRSSGSIASNYVQ 13
        |||||
Db       23 TRSSGSIASNYVQ 35
```

```
RESULT 4
US-10-269-805-50
/ Sequence 50, Application US/10269805
/ Publication No. US20030124129A1
/ GENERAL INFORMATION:
/ APPLICANT: OLINER, JONATHAN D.
/ TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
/ FILE REFERENCE: A-722
```

```
/ CURRENT APPLICATION NUMBER: US/10/269,805
/ CURRENT FILING DATE: 2002-10-10
/ PRIOR APPLICATION NUMBER: US 60/328,604
/ PRIOR FILING DATE: 2001-10-11
/ NUMBER OF SEQ ID NOS: 76
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 50
/ LENGTH: 110
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-269-805-50
```

```
Query Match          100.0%; Score 62; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TRSSGSIASNYVQ 13
        |||||
Db       23 TRSSGSIASNYVQ 35
```

```
RESULT 5
US-10-269-805-52
/ Sequence 52, Application US/10269805
/ Publication No. US20030124129A1
/ GENERAL INFORMATION:
/ APPLICANT: OLINER, JONATHAN D.
/ TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
/ FILE REFERENCE: A-722
/ CURRENT APPLICATION NUMBER: US/10/269,805
/ CURRENT FILING DATE: 2002-10-10
/ PRIOR APPLICATION NUMBER: US 60/328,604
/ PRIOR FILING DATE: 2001-10-11
/ NUMBER OF SEQ ID NOS: 76
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 52
/ LENGTH: 110
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-269-805-52
```

```
Query Match          100.0%; Score 62; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TRSSGSIASNYVQ 13
        |||||
Db       23 TRSSGSIASNYVQ 35
```

```
RESULT 6
US-10-106-698-6785
/ Sequence 6785, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptid
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: PatentIn Ver. 3.0
/ SEQ ID NO 6785
/ LENGTH: 161
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
```

NAME/KEY: MISC FEATURE
LOCATION: (149)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6785

Query Match 100.0%; Score 62; DB 15; Length 161;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|||
Db 44 TRSSGSIASNYVQ 56

RESULT 7
US-10-269-805-42
Sequence 42, Application US/10269805
Publication No. US20030124129A1
GENERAL INFORMATION:
APPLICANT: OLINER, JONATHAN D.
TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
FILE REFERENCE: A-722
CURRENT APPLICATION NUMBER: US/10/269,805
CURRENT FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: US 60/328,604
PRIOR FILING DATE: 2001-10-11
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-10-269-805-42

Query Match 95.2%; Score 59; DB 15; Length 110;
Best Local Similarity 92.3%; Pred. No. 0.0044;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|||
Db 23 TRSSGSIASNYVQ 35

RESULT 8
US-10-120-414-78
Sequence 78, Application US/10120414
Publication No. US20030175274A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig, et al.
TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
FILE REFERENCE: PF112P9
CURRENT APPLICATION NUMBER: US/10/120,414
CURRENT FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/283,385
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/250,366
PRIOR FILING DATE: 2002-01-24
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.1
SEQ ID NO 78
LENGTH: 248
TYPE: PRT
ORGANISM: Homo sapiens
US-10-120-414-78

Query Match 93.5%; Score 58; DB 12; Length 248;
Best Local Similarity 92.3%; Pred. No. 0.016;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|||
Db 160 TRSSGSIASNYVQ 172

RESULT 9
US-09-880-748-1234
Sequence 1234, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1234
LENGTH: 258
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-1234

Query Match 90.3%; Score 56; DB 11; Length 258;
Best Local Similarity 92.3%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|||
Db 170 TRSSGSIDSNYVQ 182

RESULT 10
US-09-972-656-2
Sequence 2, Application US/09972656
Publication No. US20030099647A1
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 13
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-2

Query Match 88.7%; Score 55; DB 11; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.0024;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|||
Db 1 TGSSGSIASNYVQ 13

RESULT 11
US-09-972-656-108
Sequence 108, Application US/09972656
Publication No. US20030099647A1
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra

APPLICANT: Teal, Mei-Mei
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 108
LENGTH: 216
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-108

Query Match 88.7%; Score 55; DB 11; Length 216;
Best Local Similarity 92.3%; Pred. No. 0.048;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
Db 23 TRSSGSIASNYVQ 35

RESULT 12
US-09-972-656-3
Sequence 3, Application US/09972656
Publication No. US20030099647A1
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
APPLICANT: Teal, Mei-Mei
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 13
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-3

Query Match 87.1%; Score 54; DB 11; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.0036;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
Db 1 TRSSGSIASNYVQ 13

RESULT 13
US-09-972-656-90
Sequence 90, Application US/09972656
Publication No. US20030099647A1
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
APPLICANT: Teal, Mei-Mei
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 90
LENGTH: 216
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-90

Query Match 87.1%; Score 54; DB 11; Length 216;

Best Local Similarity 92.3%; Pred. No. 0.072;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
Db 23 TRSSGSIASNYVQ 35

RESULT 14
US-09-880-748-909
Sequence 909, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 909
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-909

Query Match 85.5%; Score 53; DB 11; Length 253;
Best Local Similarity 84.6%; Pred. No. 0.13;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
Db 164 TRSSGSIASNYVQ 176

RESULT 15
US-09-880-748-1125
Sequence 1125, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1125
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-1125

Query Match 85.5%; Score 53; DB 11; Length 253;
 Best Local Similarity 84.6%; Pred. No. 0.13;
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNTYQ 13
 |||||:
 Db 164 TRSSGSIASNTYQ 176

Search completed: January 12, 2004, 07:19:56
 Job time : 28 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 07:02:25 ; Search time 34 Seconds
(without alignments)
100.755 Million cell updates/sec

Title: US-09-829-495-62

Perfect score: 89

Sequence: 1 NIKDGSSEKRYADSVRG 17

Scoring table: BLOSUM62

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	9	US-09-832-312-62
2	89	100.0	17	12	US-09-829-495-62
3	82	92.1	17	12	US-10-173-551-16
4	82	92.1	98	12	US-10-308-817-55
5	82	92.1	98	12	US-10-029-988B-85
6	82	92.1	98	12	US-10-032-037B-85
7	82	92.1	98	12	US-10-032-423A-85
8	82	92.1	98	15	US-10-194-975-15
9	82	92.1	118	12	US-10-173-551-14
10	82	92.1	120	11	US-09-995-529-8
11	82	92.1	134	9	US-09-811-737-3
12	82	92.1	241	11	US-09-880-748-1889
13	82	92.1	247	11	US-09-880-748-1470
14	82	92.1	247	12	US-10-322-673-48
15	82	92.1	251	11	US-09-880-748-1594

16	82	92.1	253	11	US-09-880-748-1003	Sequence 1003, Ap
17	82	92.1	253	11	US-09-880-748-1007	Sequence 1007, Ap
18	82	92.1	262	9	US-09-811-737-19	Sequence 19, Appl
19	82	92.1	312	12	US-10-288-917-10	Sequence 10, Appl
20	82	92.1	312	14	US-10-052-798-10	Sequence 98, Appl
21	75	84.3	17	11	US-09-977-797A-98	Sequence 1319, Ap
22	75	84.3	250	11	US-09-880-748-1319	Sequence 1486, Ap
23	75	84.3	250	11	US-09-880-748-1486	Sequence 1511, Ap
24	74	83.1	254	11	US-09-880-748-1511	Sequence 110, Appl
25	73	82.0	17	11	US-09-977-797A-110	Sequence 1350, Ap
26	73	82.0	254	11	US-09-880-748-1350	Sequence 49, Appl
27	69	77.5	124	12	US-10-010-729-49	Sequence 141, Appl
28	67	75.3	117	11	US-09-848-798-141	Sequence 1817, Ap
29	67	75.3	249	11	US-09-880-748-1817	Sequence 1909, Ap
30	66	74.2	248	11	US-09-880-748-1904	Sequence 2051, Ap
31	66	74.2	251	11	US-09-880-748-952	Sequence 914, Appl
32	66	74.2	251	11	US-09-880-748-982	Sequence 916, Appl
33	65	73.0	126	10	US-09-850-165-103	Sequence 103, Appl
34	65	73.0	241	11	US-09-880-748-1932	Sequence 1932, Ap
35	65	73.0	241	11	US-09-880-748-2054	Sequence 2054, Ap
36	65	73.0	243	11	US-09-880-748-1909	Sequence 1916, Ap
37	65	73.0	243	11	US-09-880-748-1916	Sequence 2051, Ap
38	65	73.0	243	11	US-09-880-748-2051	Sequence 2057, Ap
39	65	73.0	243	11	US-09-880-748-2057	Sequence 914, Appl
40	65	73.0	248	11	US-09-880-748-914	Sequence 916, Appl
41	65	73.0	248	11	US-09-880-748-916	Sequence 1323, Appl
42	65	73.0	249	11	US-09-880-748-1323	Sequence 2065, Ap
43	65	73.0	251	11	US-09-880-748-2065	Sequence 1114, Ap
44	65	73.0	251	11	US-09-880-748-1114	Sequence 1357, Ap
45	65	73.0	251	11	US-09-880-748-1357	

ALIGNMENTS

RESULT 1
US-09-832-312-62
; Sequence 62, Application US/09832312
; Patent No. US20010049829A1
; GENERAL INFORMATION:
; APPLICANT: Busfield et al
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-62

Query Match 100.0%; Score 89; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIKDGSSEKRYADSVRG 17
|||||
Db 1 NIKDGSSEKRYADSVRG 17

RESULT 2
US-09-829-495-62
; Sequence 62, Application US/09829495

```
/ Publication No. US20040001826A1
/ GENERAL INFORMATION:
/ APPLICANT: Busefield SJ
/ APPLICANT: Valleeval U
/ APPLICANT: Jandriot-Petrus M
/ APPLICANT: Valinchenker W
/ APPLICANT: Gili DS
/ APPLICANT: Olan MD
/ TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
/ FILE REFERENCE: 7853-234
/ CURRENT APPLICATION NUMBER: US/09/829,495
/ CURRENT FILING DATE: 2001-04-09
/ PRIOR APPLICATION NUMBER: 09/610,118
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: 09/503,387
/ PRIOR FILING DATE: 2000-02-14
/ PRIOR APPLICATION NUMBER: 09/454,824
/ PRIOR FILING DATE: 1999-12-06
/ PRIOR APPLICATION NUMBER: 09/345,468
/ PRIOR FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 62
/ LENGTH: 17
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-829-495-62
```

```
Query Match          100.0%; Score 89; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 NIKODSEKXYADSVRG 17
      |||||
DB      1 NIKODSEKXYADSVRG 17
```

```
RESULT 3
US-10-173-551-16
/ Sequence 16, Application US/10173551
/ Publication No. US2003023387A1
/ GENERAL INFORMATION:
/ APPLICANT: Lu, Chafen
/ TITLE OF INVENTION: Antibodies that bind alphaB Integrin
/ FILE REFERENCE: 1855,2025-000
/ CURRENT APPLICATION NUMBER: US/10/173,551
/ CURRENT FILING DATE: 2002-06-14
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 16
/ LENGTH: 17
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-173-551-16
```

```
Query Match          92.1%; Score 82; DB 12; Length 17;
Best Local Similarity 88.2%; Pred. No. 1.1e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 NIKODSEKXYADSVRG 17
      |||||
DB      1 NIKODSEKXYADSVRG 17
```

```
RESULT 4
US-10-308-817-55
/ Sequence 55, Application US/10308817
/ Publication No. US20030219861A1
/ GENERAL INFORMATION:
/ APPLICANT: Rother, Russell
/ APPLICANT: Wu, Dayang
/ TITLE OF INVENTION: HYBRID ANTIBODIES
/ FILE REFERENCE: 1087-37
```

```
/ CURRENT APPLICATION NUMBER: US/10/308,817
/ CURRENT FILING DATE: 2002-12-03
/ NUMBER OF SEQ ID NOS: 195
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 55
/ LENGTH: 98
/ TYPE: PRT
/ ORGANISM: human
US-10-308-817-55
```

```
Query Match          92.1%; Score 82; DB 12; Length 98;
Best Local Similarity 88.2%; Pred. No. 7.6e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 NIKODSEKXYADSVRG 17
      |||||
DB      50 NIKODSEKXYADSVRG 66
```

```
RESULT 5
US-10-029-988B-85
/ Sequence 85, Application US/10029988B
/ Publication No. US20040001839A1
/ GENERAL INFORMATION:
/ APPLICANT: Bio-Technology General Corp.
/ TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
/ FILE REFERENCE: 10793/46
/ CURRENT APPLICATION NUMBER: US/10/029,988B
/ CURRENT FILING DATE: 2001-12-31
/ PRIOR APPLICATION NUMBER: 60/258,948
/ PRIOR FILING DATE: 2000-12-29
/ NUMBER OF SEQ ID NOS: 204
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 85
/ LENGTH: 98
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-029-988B-85
```

```
Query Match          92.1%; Score 82; DB 12; Length 98;
Best Local Similarity 88.2%; Pred. No. 7.6e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 NIKODSEKXYADSVRG 17
      |||||
DB      50 NIKODSEKXYADSVRG 66
```

```
RESULT 6
US-10-032-037B-85
/ Sequence 85, Application US/10032037B
/ Publication No. US20040001822A1
/ GENERAL INFORMATION:
/ APPLICANT: Bio-Technology General Corp.
/ TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
/ FILE REFERENCE: 10793/44
/ CURRENT APPLICATION NUMBER: US/10/032,037B
/ CURRENT FILING DATE: 2001-12-31
/ PRIOR APPLICATION NUMBER: 60/258,948
/ PRIOR FILING DATE: 2000-12-29
/ NUMBER OF SEQ ID NOS: 204
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 85
/ LENGTH: 98
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-032-037B-85
```

```
Query Match          92.1%; Score 82; DB 12; Length 98;
Best Local Similarity 88.2%; Pred. No. 7.6e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```


SEQ ID NO 3
LENGTH: 134
TYPE: PRT
ORGANISM: Homo sapiens
US-09-811-737-3

Query Match 92.1%; Score 82; DB 9; Length 134;
Best Local Similarity 88.2%; Pred. No. 1,1e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYYADSVRG 17
|||||
Db 50 NIKDGSSEKYYVDSVKG 66

RESULT 12
US-09-880-748-1889
Sequence 1889, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1889
LENGTH: 241
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-1889

Query Match 92.1%; Score 82; DB 11; Length 241;
Best Local Similarity 88.2%; Pred. No. 2e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYYADSVRG 17
|||||
Db 50 NIKDGSSEKYYVDSVKG 66

RESULT 13
US-09-880-748-1470
Sequence 1470, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1470
LENGTH: 247
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-1470

Query Match 92.1%; Score 82; DB 11; Length 247;
Best Local Similarity 88.2%; Pred. No. 2,1e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYYADSVRG 17
|||||
Db 50 NIKDGSSEKYYVDSVKG 66

RESULT 14
US-10-322-673-48
Sequence 48, Application US/10322673
Publication No. US20030180296A1
GENERAL INFORMATION:

APPLICANT: Salcedo et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind to TRAIL
FILE REFERENCE: PF585
CURRENT APPLICATION NUMBER: US/10/322,673
CURRENT FILING DATE: 2002-12-19
PRIOR APPLICATION NUMBER: 60/341,237
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/369,877
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/384,828
PRIOR FILING DATE: 2002-06-04
PRIOR APPLICATION NUMBER: 60/396,591
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/403,370
PRIOR FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: 60/425,737
PRIOR FILING DATE: 2002-11-13
NUMBER OF SEQ ID NOS: 72
SEQ ID NO 48
LENGTH: 247
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: CMO13F04 scFv
US-10-322-673-48

Query Match 92.1%; Score 82; DB 12; Length 247;
Best Local Similarity 88.2%; Pred. No. 2,1e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYYADSVRG 17
|||||
Db 50 NIKDGSSEKYYVDSVKG 66

RESULT 15
US-09-880-748-1594
Sequence 1594, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 3239
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 1594
 ; LENGTH: 251
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-748-1594

Query Match 92.1%; Score 82; DB 11; Length 251;
 Best Local Similarity 88.2%; Pred. No. 2.1e-05;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKRYVDSVRG 17
 Db 50 NIKDGSSEKRYVDSVRG 66

Search completed: January 12, 2004, 07:19:52
 Job time : 34 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2004, 06:56:19 ; Search time 4.14062 seconds
(without alignments)
51.092 Million cell updates/sec

Title: US-09-829-495-61

Perfect score: 31

Sequence: 1 SYMS 5

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	5	4	US-09-832-312-61
2	31	100.0	98	1	US-07-942-245-37
3	31	100.0	117	3	US-08-545-809A-95
4	31	100.0	312	4	US-09-079-029-10
5	28	90.3	5	2	US-08-318-157B-23
6	28	90.3	10	1	US-08-208-886C-83
7	28	90.3	10	1	US-08-208-886C-87
8	28	90.3	10	1	US-08-704-744-85
9	28	90.3	10	1	US-08-704-744-89
10	28	90.3	10	1	US-08-469-557-64
11	28	90.3	10	1	US-08-469-557-68
12	28	90.3	10	1	US-08-290-793B-64
13	28	90.3	10	2	US-08-290-793B-68
14	28	90.3	118	3	US-08-767-128-2
15	28	90.3	118	3	US-08-767-128-8
16	28	90.3	119	2	US-08-318-157B-2
17	28	90.3	119	2	US-08-318-157B-8
18	28	90.3	119	2	US-08-318-157B-9
19	28	90.3	119	2	US-08-318-157B-10
20	28	90.3	119	2	US-08-318-157B-11
21	28	90.3	119	2	US-08-318-157B-12
22	28	90.3	119	2	US-08-318-157B-13
23	28	90.3	119	2	US-08-318-157B-14
24	28	90.3	119	2	US-08-318-157B-15
25	28	90.3	119	2	US-08-318-157B-17
26	28	90.3	119	2	US-08-318-157B-57
27	28	90.3	119	2	US-08-318-157B-58

28	28	90.3	119	3	US-08-767-128-6	Sequence 6, Appli
29	28	90.3	211	4	US-09-512-563C-46	Sequence 46, Appli
30	28	90.3	218	2	US-08-399-889-25	Sequence 25, Appli
31	28	90.3	218	3	US-09-167-364-25	Sequence 25, Appli
32	28	90.3	218	3	US-09-439-897-4	Sequence 4, Appli
33	28	90.3	241	1	US-08-235-838-11	Sequence 11, Appli
34	28	90.3	241	2	US-08-465-473B-11	Sequence 11, Appli
35	28	90.3	258	4	US-09-589-927-4	Sequence 4, Appli
36	28	90.3	258	4	US-09-589-927-4	Sequence 4, Appli
37	28	90.3	258	4	US-09-589-927-4	Sequence 4, Appli
38	28	90.3	260	4	US-09-589-927-2	Sequence 2, Appli
39	28	90.3	260	4	US-09-589-927-12	Sequence 12, Appli
40	28	90.3	260	4	US-09-277-665-2	Sequence 12, Appli
41	28	90.3	260	4	US-09-277-665-12	Sequence 12, Appli
42	28	90.3	260	4	US-09-589-987-2	Sequence 2, Appli
43	28	90.3	260	4	US-09-589-987-12	Sequence 12, Appli
44	28	90.3	264	4	US-09-589-927-10	Sequence 10, Appli
45	28	90.3	264	4	US-09-277-665-10	Sequence 10, Appli

ALIGNMENTS

```

RESULT 1
US-09-832-312-61
Sequence 61, Application US/09832312
Patent No. 6548741
GENERAL INFORMATION:
APPLICANT: Buefield et al
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7835-234
CURRENT APPLICATION NUMBER: US/09/832-312
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 09/610,118
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/503,387
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 09/454,824
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 09/345,468
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 61
LENGTH: 5
TYPE: PRT
ORGANISM: Homo sapiens
US-09-832-312-61

Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
Db 1 SYMS 5

RESULT 2
US-07-942-245-37
Sequence 37, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUIDI, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESS: Sugarline, Mion, Zinn, Macpeak & Seas

```

STREET: 2100 Pennsylvania Avenue, N.W.
City: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: in house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-37

Query Match 100.0%; Score 31; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
Db 31 SYMMS 35

RESULT 3
US-08-545-809A-95
Sequence 95, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tetsuko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastrSO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-95

Query Match 100.0%; Score 31; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
Db 50 SYMMS 54

RESULT 4
US-09-079-029-10
Sequence 10, Application US/09079029
Patent No. 6342369
GENERAL INFORMATION:
APPLICANT: Adams, Camilla W.
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Chuntarapai, Anan
APPLICANT: Kim, Kyung U.
TITLE OF INVENTION: Apo-2 Receptor
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,029
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P1101R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5416
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-079-029-10

Query Match 100.0%; Score 31; DB 4; Length 312;
Best Local Similarity 100.0%; Pred. No. 2,7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
Db 70 SYMMS 74

RESULT 5
US-08-318-157B-23
Sequence 23, Application US/08318157B
Patent No. 5874540
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
APPLICANT: ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED

TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
 NUMBER OF SEQUENCES: 58
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/318,157B
 FILING DATE: 05-OCT-1994
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Saxe, Bernhard D.
 REGISTRATION NUMBER: 28,665
 REFERENCE/DOCKET NUMBER: 18733/464
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 STRANDEDNESS:
 MOLECULE TYPE: protein
 US-08-318-157B-23

Query Match 90.3%; Score 28; DB 2; Length 5;
 Best Local Similarity 80.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 Db 1 TYWMS 5

RESULT 6
 US-08-208-886C-83
 Sequence 83, Application US/08208886C
 Patent No. 5597710
 GENERAL INFORMATION:
 APPLICANT: Dille, Barbara
 APPLICANT: Murgolo, Nicholas
 APPLICANT: Tindall, Stephen
 TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4
 NUMBER OF SEQUENCES: 88
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Schering-Plough Corporation
 STREET: 2000 Galloping Hill Road
 CITY: Kenilworth
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07033-0530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: Macintosh 7.1
 SOFTWARE: Microsoft Word 5.1a
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/208,886C
 FILING DATE: March 10, 1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:

FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Foulke, Cynthia L.
 REGISTRATION NUMBER: 32,364
 REFERENCE/DOCKET NUMBER: JB0429
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 908 298 2987
 TELEFAX: 908 298 5388
 INFORMATION FOR SEQ ID NO: 83:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-208-886C-83

Query Match 90.3%; Score 28; DB 1; Length 10;
 Best Local Similarity 80.0%; Pred. No. 39;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 Db 6 SYWMT 10

RESULT 7
 US-08-208-886C-87
 Sequence 87, Application US/08208886C
 Patent No. 5597710
 GENERAL INFORMATION:
 APPLICANT: Dille, Barbara
 APPLICANT: Murgolo, Nicholas
 APPLICANT: Tindall, Stephen
 TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4
 NUMBER OF SEQUENCES: 88
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Schering-Plough Corporation
 STREET: 2000 Galloping Hill Road
 CITY: Kenilworth
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07033-0530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: Macintosh 7.1
 SOFTWARE: Microsoft Word 5.1a
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/208,886C
 FILING DATE: March 10, 1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Foulke, Cynthia L.
 REGISTRATION NUMBER: 32,364
 REFERENCE/DOCKET NUMBER: JB0429
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 908 298 2987
 TELEFAX: 908 298 5388
 INFORMATION FOR SEQ ID NO: 87:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-208-886C-87

QY 1 SYMS 5
|||:
Db 6 SYMT 10

RESULT 8

US-08-704-744-85
; Sequence 85, Application US/08704744
; Patent No. 5705154
; GENERAL INFORMATION:
; APPLICANT: Daille, Barbara
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Tindall, Stephen
; TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5.3
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,744
; FILING DATE: 06-SEPT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/208886
; FILING DATE: 10-MAR-1994
; APPLICATION NUMBER: PCT/US/95/02400
; FILING DATE: 08-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Foulke, Cynthia L.
; REGISTRATION NUMBER: 32,364
; REFERENCE/DOCKET NUMBER: JB0429K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 298-2987
; TELEFAX: (908) 298-5388
; TELEX:
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-704-744-85
Query Match 90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
|||:
Db 6 SYMT 10

RESULT 9

US-08-704-744-89
; Sequence 89, Application US/08704744
; Patent No. 5705154
; GENERAL INFORMATION:
; APPLICANT: Daille, Barbara
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Tindall, Stephen
; TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4

NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.3
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/704,744
FILING DATE: 06-SEPT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208886
FILING DATE: 10-MAR-1994
APPLICATION NUMBER: PCT/US/95/02400
FILING DATE: 08-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: JB0429K
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 298-2987
TELEFAX: (908) 298-5388
TELEX:
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-704-744-89

QY 1 SYMS 5
|||:
Db 6 SYMT 10

Query Match 90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

US-08-469-557-64
; Sequence 64, Application US/08469557
; Patent No. 5770403
; GENERAL INFORMATION:
; APPLICANT: Daille, Barbara
; APPLICANT: Le, Hung
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Nguyen, Hanh
; APPLICANT: Tindall, Stephen
; APPLICANT: Zavodny, Paul
; TITLE OF INVENTION: Cloning and Expression of
; TITLE OF INVENTION: Humanized Monoclonal Antibodies
; TITLE OF INVENTION: Against Human Interleukin-4
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh

```
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,557
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,793
FILING DATE: August 16, 1994
APPLICATION NUMBER: PCT/US93/01301
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/841,659
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/782,784
FILING DATE: 24-OCT-1991
APPLICATION NUMBER: US 07/499,327
FILING DATE: 21-MAY-1990
APPLICATION NUMBER: PCT/US88/03631
FILING DATE: 21-OCT-1988
APPLICATION NUMBER: US 07/655,966
FILING DATE: 14-FEB-1991
APPLICATION NUMBER: US 07/113,623
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: US 06/881,553
FILING DATE: 03-JUL-1986
APPLICATION NUMBER: US 06/843,958
FILING DATE: 25-MAR-1986
APPLICATION NUMBER: US 06/799,668
FILING DATE: 19-NOV-1985
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: 2409K7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908 298-2987
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-557-64

Query Match      90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYMS 5
      ||||:
Db      6 SYMT 10

RESULT 11
US-08-469-557-68
; Sequence 68, Application US/08469557
; Patent No. 5770403
; GENERAL INFORMATION:
; APPLICANT: Daille, Barbara
; APPLICANT: Le, Hung
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Nguyen, Hanh
; APPLICANT: Tindall, Stephen
; APPLICANT: Zavodny, Paul
; TITLE OF INVENTION: Cloning and Expression of
; TITLE OF INVENTION: Humanized Monoclonal Antibodies
; TITLE OF INVENTION: Against Human Interleukin-4
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
```

```
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,557
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,793
FILING DATE: August 16, 1994
APPLICATION NUMBER: PCT/US93/01301
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/841,659
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/782,784
FILING DATE: 24-OCT-1991
APPLICATION NUMBER: US 07/499,327
FILING DATE: 21-MAY-1990
APPLICATION NUMBER: PCT/US88/03631
FILING DATE: 21-OCT-1988
APPLICATION NUMBER: US 07/655,966
FILING DATE: 14-FEB-1991
APPLICATION NUMBER: US 07/113,623
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: US 06/881,553
FILING DATE: 03-JUL-1986
APPLICATION NUMBER: US 06/843,958
FILING DATE: 25-MAR-1986
APPLICATION NUMBER: US 06/799,668
FILING DATE: 19-NOV-1985
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: 2409K7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908 298-2987
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-557-68

Query Match      90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYMS 5
      ||||:
Db      6 SYMT 10

RESULT 12
US-08-290-793B-64
; Sequence 64, Application US/08290793B
; Patent No. 5663537
; GENERAL INFORMATION:
; APPLICANT: Daille, Barbara
; APPLICANT: Le, Hung
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Nguyen, Hanh
; APPLICANT: Tindall, Stephen
; APPLICANT: Zavodny, Paul
; TITLE OF INVENTION: Cloning and Expression of
; TITLE OF INVENTION: Humanized Monoclonal Antibodies
```

;; TITLE OF INVENTION: Against Human Interleukin-4
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corporation
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07033-0530
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Apple Macintosh
;; OPERATING SYSTEM: Macintosh 6.0.5
;; SOFTWARE: Microsoft Word 5.1A
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/290,793B
;; FILING DATE: August 16, 1994
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/01301
;; FILING DATE: 19-FEB-1992
;; APPLICATION NUMBER: US 07/841,659
;; FILING DATE: 19-FEB-1992
;; APPLICATION NUMBER: US 07/782,784
;; FILING DATE: 24-OCT-1991
;; APPLICATION NUMBER: US 07/499,327
;; FILING DATE: 21-MAY-1990
;; APPLICATION NUMBER: PCT/US88/03631
;; FILING DATE: 21-OCT-1988
;; APPLICATION NUMBER: US 07/655,966
;; FILING DATE: 14-FEB-1991
;; APPLICATION NUMBER: US 07/113,623
;; FILING DATE: 26-OCT-1987
;; APPLICATION NUMBER: US 06/881,553
;; FILING DATE: 03-JUL-1986
;; APPLICATION NUMBER: US 06/843,958
;; FILING DATE: 25-MAR-1986
;; APPLICATION NUMBER: US 06/799,668
;; FILING DATE: 19-NOV-1985
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Foulke, Cynthia L.
;; REGISTRATION NUMBER: 32,364
;; REFERENCE/DOCKET NUMBER: 2409K7
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908 298-2987
;; TELEFAX: 908-298-5388
;; INFORMATION FOR SEQ ID NO: 64:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;;
;; US-08-290-793B-64
;;
Query Match 90.3%; Score 28; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
| | | | |
Db 6 SYMT 10

RESULT 13
US-08-290-793B-68
; Sequence 68, Application US/08290793B
; Patent No. 5863537
; GENERAL INFORMATION:
; APPLICANT: Dalle, Barbara
; APPLICANT: Le, Hung
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Nguyen, Hanh
; APPLICANT: Tindall, Stephen

;; APPLICANT: Zavodny, Paul
;; TITLE OF INVENTION: Cloning and Expression of
;; TITLE OF INVENTION: Humanized Monoclonal Antibodies
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corporation
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07033-0530
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Apple Macintosh
;; OPERATING SYSTEM: Macintosh 6.0.5
;; SOFTWARE: Microsoft Word 5.1A
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/290,793B
;; FILING DATE: August 16, 1994
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/01301
;; FILING DATE: 19-FEB-1992
;; APPLICATION NUMBER: US 07/841,659
;; FILING DATE: 19-FEB-1992
;; APPLICATION NUMBER: US 07/782,784
;; FILING DATE: 24-OCT-1991
;; APPLICATION NUMBER: US 07/499,327
;; FILING DATE: 21-MAY-1990
;; APPLICATION NUMBER: PCT/US88/03631
;; FILING DATE: 21-OCT-1988
;; APPLICATION NUMBER: US 07/655,966
;; FILING DATE: 14-FEB-1991
;; APPLICATION NUMBER: US 07/113,623
;; FILING DATE: 26-OCT-1987
;; APPLICATION NUMBER: US 06/881,553
;; FILING DATE: 03-JUL-1986
;; APPLICATION NUMBER: US 06/843,958
;; FILING DATE: 25-MAR-1986
;; APPLICATION NUMBER: US 06/799,668
;; FILING DATE: 19-NOV-1985
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Foulke, Cynthia L.
;; REGISTRATION NUMBER: 32,364
;; REFERENCE/DOCKET NUMBER: 2409K7
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908 298-2987
;; TELEFAX: 908-298-5388
;; INFORMATION FOR SEQ ID NO: 68:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;;
;; US-08-290-793B-68
;;
Query Match 90.3%; Score 28; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
| | | | |
Db 6 SYMT 10

RESULT 14
US-08-767-128-2
; Sequence 2, Application US/08767128
; Patent No. 611079
; GENERAL INFORMATION:
; APPLICANT: WYLE, DWANE E.
; APPLICANT: LOPEZ, OSVALDO
; APPLICANT: MURRAY, PETER JOSEPH

APPLICANT: GOEBEL, PETER
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND
TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSES: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 611079west Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,128
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 04-DEC-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09258
FILING DATE: 05-JUN-1996
PRIOR APPLICATION DATA: 08/541,373
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,798
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648.49USF1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/371-5278
TELEFAX: 612/332-9081
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-767-128-2

Query Match 90.3%; Score 28; DB 3; Length 118;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
Db 31 SYWMN 35

RESULT 15
US-08-767-128-8
Sequence 8, Application US/08767128
Patent No. 6111079
GENERAL INFORMATION:
APPLICANT: WYLIE, DWANE E.
APPLICANT: LOPEZ, OSVALDO
APPLICANT: MURRAY, PETER JOSEPH
APPLICANT: GOEBEL, PETER
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND

TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSES: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 611079west Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,128
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 04-DEC-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09258
FILING DATE: 05-JUN-1996
PRIOR APPLICATION DATA: 08/541,373
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,798
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648.49USF1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/371-5278
TELEFAX: 612/332-9081
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-767-128-8

Query Match 90.3%; Score 28; DB 3; Length 118;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
Db 31 SYWMN 35

Search completed: January 12, 2004, 07:03:57
Job time : 6.14062 secs

THIS PAGE BLANK (USPTO)

PI Busfield SJ, Villelaj J, Jandrot-Perrus M, Vainchenker W, Gali DS,
 PI Qian DM, Kingsbury G;
 XX
 XX WPI; 2001-080877/09.
 DR
 XX New genes encoding human platelet-expressed collagen receptor,
 PT glycoprotein VI, and its modulators, useful for preventing, treating
 PT and diagnosing hemorrhagic disorders, thrombotic diseases and
 PT immunological disorders -
 XX
 PS Claim 31; Page 102; 227pp; English.
 XX
 CC The present sequence is given in a specification relating to an isolated
 CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
 CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
 CC and polypeptides and their modulators, e.g. antisense nucleic acids,
 CC ribozymes and antibodies, are useful for preventing, treating and
 CC diagnosing disorders associated with aberrant expression or activity of
 CC GPVI. These disorders include bleeding disorders
 CC (e.g. thrombocytopenia), blood vessel injury, thrombotic disorders
 CC (e.g. thrombotic occlusion of the coronary arteries), hemorrhagic
 CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
 CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
 CC infarction), immunological diseases (e.g. platelet disorder) and
 CC embryonic liver disorders. Preferably they are used to prevent accurate
 CC cardiac ischaemia following angioplasty and metastatic cancers,
 CC especially of the colon and liver.
 CC
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 31; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYMMS 5
 Db 1 SYMMS 5
 RESULT 2
 ID ABU11259 standard; Peptide; 5 AA.
 XX
 AC ABU11259;
 XX
 DT 06-FEB-2003 (first entry)
 XX
 DE Human TANGO 268 VHCPR1 Peptide #3.
 XX
 KM Human; mouse; variable heavy; VH; antigen; cancer;
 KM Complementarity determining region; TANGO 268; glycoprotein VI; GPVI;
 KM TANGO 268; extracellular matrix; collagen; platelet release;
 KM proliferation; migration; embryogenesis; inflammation; thrombosis;
 KM degranulation; thrombocytopenia; antibody; thrombotic disorder;
 KM cerebral vascular disease; stroke; ischaemia; venous thromboembolism;
 KM leg swelling; pain; ulceration; pulmonary embolism; coronary disease;
 KM cardiovascular disease; angina pectoris; myocardial infarction;
 KM coronary restenosis; atherosclerosis; immunological disorder;
 KM developmental disorder; embryonic disorder; liver disorder;
 KM cerebral vascular disease; venous thromboembolism disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200280968-A1.
 XX
 PD 17-OCT-2002.
 XX
 PF 09-APR-2002; 2002WO-US11122.
 XX
 PR 09-APR-2001; 2001US-0829495.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX

PI Busfield SJ, Villelaj J, Jandrot-Perrus M, Vainchenker W, Gali DS,
 PI Qian DM, Kingsbury G;
 XX
 XX WPI; 2003-058477/05.
 DR
 XX Novel substantially purified antibody immunospecifically binding to
 PT TANGO 268 antigen, useful for treating bleeding disorders such as
 PT thrombocytopenia, stroke, ischaemia, pulmonary embolism, atherosclerosis
 PT
 XX
 PS Disclosure; Page 111; 236pp; English.
 XX
 CC This invention relates to a novel purified antibody comprising a
 CC variable heavy (VH) complementarity determining region (CDR)1, VH CDR2
 CC or VH CDR3; or variable light (VL) CDR1, VL CDR2 or VL CDR3, and
 CC immunospecifically binding to a TANGO 268 (also referred as glycoprotein
 CC VI (GPVI)) antigen. The antibodies of the invention act to decrease or
 CC block TANGO 268 binding to extracellular matrix components, or as a
 CC Collagen or platelet release and aggregation blocker. The antibodies of
 CC the invention are useful for modulating proliferation, migration,
 CC morphology, differentiation and/or function of megakaryocytes and
 CC platelets, including during development e.g. embryogenesis, modulating
 CC leukocyte-platelet and platelet-endothelium interactions in
 CC inflammation and/or thrombosis, and modulating platelet aggregation and
 CC degranulation. They are also useful for modulating disorders associated
 CC with abnormal or aberrant megakaryocyte and/or platelet proliferation,
 CC migration, morphology, differentiation and/or function, e.g. bleeding
 CC disorders such as thrombocytopenia. Other diseases which may be
 CC modulated by these antibodies are thrombotic disorders, cerebral
 CC vascular diseases (e.g. stroke and ischaemia) venous thromboembolism
 CC diseases (e.g. diseases involving leg swelling, pain and ulceration,
 CC pulmonary embolism, etc); coronary diseases (e.g. cardiovascular
 CC diseases including angina pectoris, myocardial infarction, coronary
 CC restenosis, atherosclerosis, etc); immunological disorders,
 CC developmental disorders, embryonic disorders, liver disorders, cerebral
 CC vascular diseases, venous thromboembolism disease, coronary diseases,
 CC and metastatic cancers. The antibodies of the invention only causes a
 CC transient decrease in platelet counts, platelet aggregation, and/or
 CC platelet activation and so have some advantages over prior art
 CC methods. The present sequence represents a peptide sequence used to
 CC generate the antibodies of the invention.
 CC
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 31; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYMMS 5
 Db 1 SYMMS 5
 RESULT 3
 ID AAR52066 standard; Protein; 98 AA.
 XX
 AC AAR52066;
 XX
 DT 11-OCT-1996 (first entry)
 XX
 DE Heavy chain variable region of human PL0123 antibody.
 XX
 KM antibody; humanised; murine; human; heavy chain; light; variable;
 KM framework region; complementarity determining region; reshaping;
 KM modelling; surface residue; modify.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Region
 FT 1..30 Location/Qualifiers
 FT /label=framework_region_1
 FT /note="PR 1"

	Region	31..35	/label= complementarity_determining_region_1
FT		/note= "CDR 1"	
FT	Region	36..49	
FT		/note= "FR 2"	
FT	Region	50..59	
FT		/note= "CDR 2"	
FT	Region	60..98	
FT		/note= "FR 3"	
PX			
PN	EP592106-A1.		
PD	13-APR-1994.		
PX			
XX	07-SEP-1993;	93EP-0307051.	
XX			
PR	09-SEP-1992;	92US-0942245.	
PA	(PEDE/) PEDERSEN J T.		
PA	(IMMU-) IMMUNOGEN INC.		
P1	Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;		
DR	WPI; 1994-120230/15.		
XX			
PT	Method of resurfacing of rodent antibodies to produce humanised antibody forms - for producing non-human antibodies with improved therapeutic efficiency by presenting human surface on V-region		
PS	Example 1; Fig 4B; 23opp; English.		
XX			
CC	Modification of a rodent antibody (Ab) or fragment by resurfacing in order to produce a humanised rodent Ab can be determined by calculating homology between murine and human Ab antibody surfaces. In order to test the resurfacing approach of the invention, three humanisation experiments were set up: (1) traditional loop grafting; (2) resurfacing approach using most similar chain; and (3) resurfacing approach using human sequences with most similar surface residues. The Ab used was the murine anti-N901 Ab (see AAR52061). Experiment 3 was carried out using the present sequence which represents the human PL0123 Ab light chain variable region with 74 percent homology with anti-N901 Ab. N901/PL0123 (AAR52067) was prepd. by resurfacing. Sequence numbering starts at position 118 in the specification.		
CC			
CC			
CC			
CC			
CC			
CC			
SC	Sequence 98 AA;		
SQ			
	Query Match	100.0%; Score 31; DB 15; Length 98;	
	Best Local Similarity	100.0%; Pred. No. 1.4e+02;	
	Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 SYMS 5 		
Dd	31 SYMS 35		
	RESULT 4		
ID	AAR72074		
AC	AAR72074 standard; Protein; 98 AA.		
XX			
XX	AAR72074;		
DT	25-MAR-2003 (updated)		
DT	26-SEP-1995 (first entry)		
XX			
DE	DP54 VH region.		
XX			
KM	Graves ophthalmopathy associated immunoglobulin protein;		
KW	orbital antigen; monoclonal antibody; heavy chain; H chain; variable region; autoimmunity.		
OS	Homo sapiens.		
XX			
FH	Key Location/Qualifiers		

FT	Region	41..44
PT	/label= CDR1	
FT	Region	49..67
FT	/label= CDR2	
XX		
PX	W09508336-A1.	
XX		
PD	30-MAR-1995.	
XX		
PF	22-SEP-1994;	94WO-US10756.
PR	22-SEP-1993;	93US-0124469.
XX		
PA	(NICH-) NICHOLS INST DIAGNOSTICS.	
PI	McLachlan SM, Rapoport B;	
XX		
DR	WPI; 1995-139383/18.	
N-PSDB; AAQ89332.		
XX		
PT	Graves' ophthalmopathy-associated monoclonal antibody - produced by molecular cloning of immunoglobulin genes by PCR	
XX		
PS	Disclosure; Page 72; 94pp; English.	
CC	L- and H-chain DNA was amplified by PCR from Graves' orbital tissue and clones encoding autoimmune-associated immunoglobulin fragments were obtained. 2/15 clones of H chain (IgG1) genes showed homology to the germ-line gene DP54 (AAO89332), which encodes the protein given in AAR72074. The DNA (AAQ89333) and corresp. amino acid (AAR72075) sequences of the VH region of a representative clone, OE7H1.19, are provided.	
CC	(Updated on 25-MAR-2003 to correct PN field.)	
SO	Sequence 98 AA;	
OY	1 SYMMS 5 31 SYMMS 35	Score 31; DB 16; Length 98; Pred. No. 1.4e+02; Indels 0; Gaps 0;
ID	AAB40138 standard; Protein; 98 AA.	
AC	AAB40138;	
DT	05-FEB-2001 (first entry)	
DE	Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 664.	
XX		
KW	Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory; complementarity determining region; CDR; antineumatic; antiarthritic; antisclerotic; neuroprotective; antipsoriatic; antiaesthetic; cardiac; antiparasitic; antibacterial; immunosuppressive; Crohn's disease; multiple sclerosis; rheumatoid arthritis.	
OS	Homo sapiens.	
PN	W0200056772-A1.	
XX		
PF	28-SEP-2000.	
PR	24-MAR-2000; 2000MO-US07946.	
XX		
PR	25-MAR-1999; 99US-0126603.	
XX		
PA	(BADI) BASF AG.	
PA	(GEVY) GENETICS INST INC.	

XX Salfield JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M,
PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Vaidman GM, Venturini A, Warne NW, Widom A, Elyin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holter TL, Du Fou SL;
XX WPI; 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
PS
PS Claim 75; Page 122; 377pp; English.
XX
XX This invention relates to a new human antibody specific for human
XX interleukin-12 (IL-12). The invention also includes antigen binding
XX portions that bind to IL-12. Sequences AAB39485-B39516 represent human
XX anti-IL-12 antibody heavy and light chain complementarity determining
XX region (CDR) amino acid sequences, and also includes variable region
XX amino acid sequences. Other variable region amino acid sequences are
XX given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
XX represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063
XX represent other CDR sequences. Light chain CDR3 consensus sequences are
XX given in AAB40064-B40067. Primers used in the identification and
XX construction of the antibodies of the invention are given in
XX AAC61062-C61071. The antibody of the invention is a neutralising
XX antibody and has antineumatic; antiarthritic; antisclerotic;
XX antiinflammatory; neuroprotective; antipsoriatic; antiaesthetic;
XX cardiant; antiparasitic; antibacterial and immunosuppressive activity.
XX The antibodies or antigen-binding fragments are useful in the treatment
XX of disorders associated with detrimental release of human IL-12,
XX especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
XX They can also be used in the manufacture of a pharmaceutical composition
XX to treat human IL-12 disorders.
SQ
SQ Sequence 98 AA;
Query March 100.0%; Score 31; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYMWS 5
Db 31 SYMWS 35
RESULT 6
AAB40140
ID AAB40140 standard; Protein; 98 AA.
XX
AC AAB40140;
XX
DT 05-FEB-2001 (first entry)
XX
XX Anti-IL12 antibody H chain V region amino acid sequence SEQ ID 666.
DE
XX Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
XX complementarity determining region; CDR; antineumatic; antiarthritic;
XX antisclerotic; neuroprotective; antipsoriatic; antiaesthetic; cardiant;
XX antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
XX multiple sclerosis; rheumatoid arthritis.
OS Homo sapiens.
XX
XX WO200056772-A1.
PN
XX 28-SEP-2000.
PD
XX 24-MAR-2000; 2000WO-US07946.
PF
XX 25-MAR-1999; 99US-0126603.
PR
XX (BADI) BASF AG.
PA

PA (GBM) GENETICS INST INC.
XX Salfield JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M,
PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Vaidman GM, Venturini A, Warne NW, Widom A, Elyin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holter TL, Du Fou SL;
XX WPI; 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
PS
PS Claim 75; Page 122; 377pp; English.
XX
XX This invention relates to a new human antibody specific for human
XX interleukin-12 (IL-12). The invention also includes antigen binding
XX portions that bind to IL-12. Sequences AAB39485-B39516 represent human
XX anti-IL-12 antibody heavy and light chain complementarity determining
XX region (CDR) amino acid sequences, and also includes variable region
XX amino acid sequences. Other variable region amino acid sequences are
XX given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
XX represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063
XX represent other CDR sequences. Light chain CDR3 consensus sequences are
XX given in AAB40064-B40067. Primers used in the identification and
XX construction of the antibodies of the invention are given in
XX AAC61062-C61071. The antibody of the invention is a neutralising
XX antibody and has antineumatic; antiarthritic; antisclerotic;
XX antiinflammatory; neuroprotective; antipsoriatic; antiaesthetic;
XX cardiant; antiparasitic; antibacterial and immunosuppressive activity.
XX The antibodies or antigen-binding fragments are useful in the treatment
XX of disorders associated with detrimental release of human IL-12,
XX especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
XX They can also be used in the manufacture of a pharmaceutical composition
XX to treat human IL-12 disorders.
SQ
SQ Sequence 98 AA;
Query March 100.0%; Score 31; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYMWS 5
Db 31 SYMWS 35
RESULT 7
ABG91901
ID ABG91901 standard; Protein; 98 AA.
XX
AC ABG91901;
XX
DT 04-DEC-2002 (first entry)
XX
XX Human antibody fragment #85.
DE
XX Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
XX metastasis; hypervariable region; autoimmune disease; thrombosis;
XX leukaemia; inflammatory disease; cardiovascular disease;
XX myocardial infarction; retinopathic disease; abnormal platelet function;
XX sulphated tyrosine-dependent protein-protein interaction.
OS Homo sapiens.
XX
XX WO200253700-A2.
PN
XX 11-JUL-2002.
PD
XX 31-DEC-2001; 2001WO-US49442.
PF
XX 29-DEC-2000; 2000US-258948P.
PR
XX 29-DEC-2000; 2000US-0751181.
PA

XX	PA	(BIOT-) BIO-TECHNOLOGY GEN CORP.
XX	PI	Lazarovitch J, Hggai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H,
XX	PI	Szananton E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX	XX	WPI, 2002-674776/72.
XX	PT	Novel isolated epitope present on cancer cells and important in
XX	PT	physiological phenomena such as cell rolling, metastasis and
XX	PT	inflammation, for treating autoimmune, inflammatory or cardiovascular
XX	PS	diseases, and cancer
XX	PS	Disclosure: Page 268; 310pp; English.
XX	CC	The invention relates to an isolated epitope present on cancer cells and
XX	CC	important in physiological phenomena such as cell rolling, metastasis and
XX	CC	inflammation, where the epitope is capable of being bound by an antibody,
XX	CC	its antigen-binding fragment or its complex comprising at least one
XX	CC	antibody or its binding fragment having a first hypervariable region. The
XX	CC	epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
XX	CC	disease, thrombosis, restenosis, metastasis, growth and/or replication of
XX	CC	tumour or leukaemia cells, increase in number of tumour or leukaemia
XX	CC	cells in a patient, cell-cell, cell-matrix, platelet-matrix,
XX	CC	platelet-platelet and/or cell-platelet adhesion or aggregation, for
XX	CC	increasing mortality of diseased cells to damage by anti-disease, anti-cancer
XX	CC	or anti-leukaemia agents, or for decreasing the number of tumour or
XX	CC	leukaemia cells in a patient, or in the manufacture of a medicament for
XX	CC	the above mentioned purposes. The epitopes are useful for diagnosing and
XX	CC	treating diseases such as cancer, leukaemia, autoimmune diseases,
XX	CC	inflammatory diseases, cardiovascular diseases such as myocardial
XX	CC	infarction, retinopathic diseases and other diseases mediated by abnormal
XX	CC	platelet function and diseases caused by sulphated tyrosine-dependent
XX	CC	protein-protein interactions. This sequence represents a human antibody
XX	CC	fragment of the invention.
XX	XX	Sequence 98 AA;
XX	QY	Query Match 100.0%; Score 31; DB 23; Length 98;
XX	Db	Best Local Similarity 100.0%; Pred. No. 1.4e+02;
XX	QY	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX	QY	1 SYMMS 5
XX	QY	
XX	QY	31 SYMMS 35
XX	RESULT 8	
XX	ABG78210	
XX	ID	ABG78210 standard; Protein; 98 AA.
XX	XX	
XX	AC	ABG78210;
XX	DT	
XX	DT	15-NOV-2002 (first entry)
XX	DE	
XX	DE	Human Fv molecule hypervariable region related peptide #95.
XX	KW	
XX	KW	Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
XX	KW	disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
XX	KW	Lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX	OS	
XX	OS	Homo sapiens.
XX	PN	
XX	PN	WO200259264-A2.
XX	PD	
XX	PD	01-AUG-2002.
XX	PF	
XX	PF	31-DEC-2001; 2001WO-US49440.
XX	RR	
XX	RR	29-DEC-2000; 2000US-0751181.
XX	XX	
XX	XX	(BIOT-) BIO-TECHNOLOGY GEN CORP.

```

XX Hagai Y, Lazarovits J, Guy R, Lipechitz O, Stanton E, Levanon A;
XX P1 Plasmid D, Perez T;
XX DR WPI; 2002-619166/66.
XX
XX Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
XX or fragment, or construct of fragment with enhanced binding
XX characteristics so as to selectively bind target cell in favour of other
XX cells -
XX
XX Claim 13; Page 191-192; 232pp; English.
XX
XX The invention relates to a peptide or polypeptide comprising an Fv
XX molecule, a construct or fragments or a construct of a fragment with
XX enhanced binding characteristics which selectively and/or specifically
XX binds to a target cell in favour of other cells, where binding is
XX primarily determined by a first hypervariable region and Fv is a single
XX chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
XX association with or attached, coupled, combined, linked or fused to a
XX pharmaceutical agent, is useful in the manufacture of a medicament, where
XX the medicament has activity against a diseased cell, preferably a cancer
XX cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
XX myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
XX acute myeloid leukaemia cell). The peptide is also useful for preparing a
XX composition for use in inhibiting the growth of a diseased or cancer
XX cell. This sequence represents a human Fv molecule hypervariable region
XX related peptide of the invention.
XX
XX Sequence 98 AA;
XX
XX Query Match 100.0%; Score 31; DB 23; Length 98;
XX Best Local Similarity 100.0%; Pred. NO. 1,4e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SYMMS 5
XX |||||
XX 31 SYMMS 35
XX
XX RESULT 9
XX AAR66301
XX AAR66301 standard; Protein; 117 AA.
XX
XX AAR66301;
XX
XX AC 25-MAR-2003 (updated)
XX DT 02-AUG-1995 (first entry)
XX
XX Human immunoglobulin variable heavy chain #7.
XX
XX Primer; PCR; amplify; human; immunoglobulin; variable; heavy chain;
XX cosmid; placenta; vector; pJB81; E.Coli; mammalian.
XX
XX Homo sapiens.
XX
XX WO9426895-A1.
XX
XX 24-NOV-1994.
XX
XX 10-MAY-1993; 93WO-JP00603.
XX
XX 10-MAY-1993; 93WO-JP00603.
XX
XX (NLSB ) JAPAN TOBACCO INC.
XX
XX Honjo T, Matsuda F;
XX
XX WPI; 1995-006791/01.
XX
XX N-PSDB; AAO78946.
XX
XX DNA fragment comprising human immunoglobulin Vh genes - for the
XX production of human immunoglobulin in mammalian hosts
XX

```

XX
PS Claim 17; Page 40-41; 130pp; Japanese.

CC Protein sequences (AA66295-51) are novel human immunoglobulin heavy
CC chain sequences encoded by novel isolated genes. The genes
CC (AAQ78939-79002) were isolated and cloned from a series of cosmid
CC constructs: Y202; Y103; Y21; Y6; Y24; 3-31; M84; M18 and M31, by PCR
CC amplification using primers AAQ78917-38. The genes are subdivided into 5
CC families of Vh genes. The fragments cover a region of 800 kb. The DNA
CC fragments were isolated from high molecular weight DNA from human
CC placenta. The DNA was partially digested with *TaqI* restriction enzyme.
CC The fragments were separated by gel electrophoresis and 35-45 kb fractions
CC were collected. The fragments were ligated with ClaI-digested cosmid
CC vector pUB91. The ligation products were then subcloned by colony
CC into *E. coli* 4908. The fragments were then subcloned by colony
CC hybridization. The Vh genes and the DNA fragments encoding them are
CC useful in producing human immunoglobulin in mammalian hosts.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX
SQ Sequence 117 AA;

Query Match 100.0%; Score 31; DB 16; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
Db 50 SYMMS 54

RESULT 10
AAE25526
ID AAE25526 standard; Protein; 117 AA.

XX
AC AAE25526;

DT 04-NOV-2002 (first entry)

XX
DE S. aureus Map10 antibody H07VHD-1 (variable heavy chain) protein.

XX
KM Map10 protein; infection; immunological response; passive immunisation;
KM vaccine; antibacterial; variable heavy chain; H07VHD-1 protein.

XX
OS Staphylococcus aureus.

XX
FH Key Location/Qualifiers
FH Region 31..35
FT /note= "Complementarity determining region (CDR)"
FT Region 53..66
FT /note= "Complementarity determining region (CDR)"
FT Region 99..106
FT /note= "Complementarity determining region (CDR)"

XX
PN WO200234788-A1.

XX
PD 02-MAY-2002.

XX
PF 22-OCT-2001; 2001WO-US32550.

XX
PR 20-OCT-2000; 2000US-241832P.

XX
PR 21-MAR-2001; 2001US-277287P.

XX
PA (INH1-) INHIBITEX INC.

XX
PI Patcl JM, Domanski P, Patel P;

XX
DR WPI: 2002-547435/58.

XX
DR N-PSDB; AAD41846.

XX
PT Antibody capable of binding subdomains of Map10 protein, a surface
XX localized protein from *Staphylococcus aureus*, useful for preventing and
XX treating staphylococcal infection in humans or animals

PS Claim 11; Page 48; 52pp; English.

XX
XX The present sequence relates to novel antibodies which are capable of
CC binding to the subdomains Map10 protein, a surface localised protein
CC from *Staphylococcus aureus*. Sequences of the invention prevent *S. aureus*
CC infection in a human or animal and inhibit binding of staphylococcal
CC bacteria to eukaryotic cells. They are also useful for diagnosing an
CC infection in humans or animals. Antibodies of the invention are useful
CC for inducing an immunological response in a human or animal and for
CC identifying antibodies to the Map10 protein in a sample. They are also
CC useful in preventing adherence of *Staphylococcal* bacteria, in production
CC facilities or laboratories to isolate additional quantities of the
CC proteins, such as by affinity chromatography and in the development of
CC vaccines for passive immunisation against staphylococcal infections. The
CC antibodies prevent or reduce bacterial infection on in-dwelling medical
CC devices to make them safer to use, including sutures, replacement heart
CC valves, cardiac assist devices, hard and soft contact lenses, intraocular
CC lens implants, other implants such as corneal inlays, kerato-protheses,
CC vascular stents, dental protheses, pacemakers and heart valves. The
CC present sequence is a DNA encoding *Staphylococcus aureus* Map10 antibody
CC H07VHD-1 (variable heavy chain) protein.

XX
SQ Sequence 117 AA;

Query Match 100.0%; Score 31; DB 23; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
Db 31 SYMMS 35

RESULT 11
AAB62777
ID AAB62777 standard; Protein; 123 AA.

XX
AC AAB62777;

DT 03-APR-2001 (first entry)

XX
DE Human HIV-1 monoclonal antibody SEQ ID NO: 76.

XX
KM Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;
KM envelope glycoprotein; gp120; diagnosis.

XX
OS Homo sapiens.

XX
OS WO200100678-A1.

XX
PN 04-JAN-2001.

XX
PD 23-JUN-2000; 2000WO-US17327.

XX
PF 30-JUN-1999; 99US-0141701.

XX
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX
PA Wackins BA, Reitz MS;

XX
PI WPI: 2001-112438/12.

XX
DR N-PSDB; AAF29078.

XX
PT Novel human monoclonal antibody immunoreactive with human
XX immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
XX in biological sample and providing passive immunotherapy to HIV-1
XX infected mammal

XX
PS Claim 1; Page 70; 81pp; English.

XX
XX The present invention provides the protein and coding sequences for the
XX variable regions of human monoclonal antibodies which are immunoreactive

CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
 CC These can be used in diagnosis and therapy of HIV-1 infection.
 XX
 XX
 SQ Sequence 123 AA;

Query Match 100.0%; Score 31; DB 22; Length 123;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 |||||
 Db 32 SYWMS 36

RESULT 12
 AAB62778
 ID AAB62778 standard; Protein; 123 AA.

AC AAB62778;
 XX
 DT 03-APR-2001 (first entry)
 XX

DE Human HIV-1 monoclonal antibody SEQ ID NO: 77.

KW Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;
 KM envelope glycoprotein; gp120; diagnosis.

OS Homo sapiens.

PN WO200100678-A1;
 XX

PD 04-JAN-2001.
 XX

PF 23-JUN-2000; 2000WO-US17327.
 XX

PR 30-JUN-1999; 99US-0141701.
 XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX

PI Watkins BA, Reitz MS;
 XX

DR WPI; 2001-112438/12.
 XX

DR N-PSDB; AAF29079.
 XX

PT Novel human monoclonal antibody immunoreactive with human
 PT immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
 PT in biological sample and providing passive immunotherapy to HIV-1
 PT infected mammal -

PS Claim 1; Page 71; 81pp; English.
 XX

CC The present invention provides the protein and coding sequences for the
 CC variable regions of human monoclonal antibodies which are immunoreactive
 CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
 CC These can be used in diagnosis and therapy of HIV-1 infection.
 XX

SQ Sequence 123 AA;

Query Match 100.0%; Score 31; DB 22; Length 123;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 |||||
 Db 32 SYWMS 36

RESULT 13
 AAB62776
 ID AAB62776 standard; Protein; 125 AA.
 XX
 AC AAB62776;
 XX

DT 03-APR-2001 (first entry)
 XX
 DE Human HIV-1 monoclonal antibody SEQ ID NO: 75.
 XX

KW Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;
 KM envelope glycoprotein; gp120; diagnosis.

OS Homo sapiens.

PN WO200100678-A1.
 XX

PD 04-JAN-2001.
 XX

PF 23-JUN-2000; 2000WO-US17327.
 XX

PR 30-JUN-1999; 99US-0141701.
 XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX

PI Watkins BA, Reitz MS;
 XX

DR WPI; 2001-112438/12.
 XX

DR N-PSDB; AAF29077.
 XX

PT Novel human monoclonal antibody immunoreactive with human
 PT immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
 PT in biological sample and providing passive immunotherapy to HIV-1
 PT infected mammal -

PS Claim 1; Page 69-70; 81pp; English.
 XX

CC The present invention provides the protein and coding sequences for the
 CC variable regions of human monoclonal antibodies which are immunoreactive
 CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
 CC These can be used in diagnosis and therapy of HIV-1 infection.
 XX

SQ Sequence 125 AA;

Query Match 100.0%; Score 31; DB 22; Length 125;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 |||||
 Db 32 SYWMS 36

RESULT 14
 ABP45878

ID ABP45878 standard; Protein; 241 AA.
 XX

AC ABP45878;
 XX

DT 19-AUG-2002 (first entry)
 XX

DE Human Bly's binding scFv. SEQ ID 1889.
 XX

KW Bly's B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KM tumour necrosis factor; B cell proliferation; B cell differentiation;
 KM immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KM antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KM systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KM common variable immunodeficiency; acquired immunodeficiency syndrome.

OS Homo sapiens.
 XX

PN WO200202641-A1.
 XX

PD 10-JAN-2002.
 XX

PF 15-JUN-2001; 2001WO-US19110.
 XX

PR 16-JUN-2000; 2000US-212210P.
 XX

PR 17-OCT-2000; 2000US-240816P.
 PR 16-MAR-2001; 2001US-276248P.
 PR 21-MAR-2001; 2001US-277379P.
 PR 25-MAY-2001; 2001US-293499P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 PI WPI; 2002-114799/15.
 DR
 XX
 PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for
 PT the diagnosis and treatment of cancers and immune disorders -
 XX
 PS Claim 1; Page 2655-2656; 3148pp; English.
 XX
 CC This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method
 CC of the invention.
 CC
 SQ Sequence 241 AA;
 XX
 XX
 Query Match 100.0%; Score 31; DB 23; Length 241;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYMMS 5
 |||||
 Db 31 SYMMS 35
 XX
 RESULT 15
 ABP45459
 ID ABP45459 standard; Protein; 247 AA.
 AC ABP45459;
 XX
 XX 19-AUG-2002 (first entry)
 DT
 XX
 DE Human Blys binding scfv SEQ ID 1470.
 XX
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO200202641-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 15-JUN-2001; 2001WO-US19110.
 XX
 XX 16-JUN-2000; 2000US-212210P.
 PR 17-OCT-2000; 2000US-240816P.

PR 16-MAR-2001; 2001US-276248P.
 PR 21-MAR-2001; 2001US-277379P.
 PR 25-MAY-2001; 2001US-293499P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 PI WPI; 2002-114799/15.
 DR
 XX
 PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for
 PT the diagnosis and treatment of cancers and immune disorders -
 XX
 PS Claim 1; Page 2157-2158; 3148pp; English.
 XX
 CC This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method
 CC of the invention.
 CC
 SQ Sequence 247 AA;
 XX
 XX
 Query Match 100.0%; Score 31; DB 23; Length 247;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYMMS 5
 |||||
 Db 31 SYMMS 35
 XX
 Search completed: January 12, 2004, 06:58:52
 Job time : 14.1094 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 07:02:25 ; Search time 10 Seconds
(without alignments)
100.755 Million cell updates/sec

Title: US-09-829-495-61

Perfect score: 31
Sequence: 1 SYMS 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	5	US-09-832-312-61	Sequence 61, Appl
2	31	100.0	5	US-09-829-495-61	Sequence 55, Appl
3	31	100.0	98	US-10-308-817-55	Sequence 85, Appl
4	31	100.0	98	US-10-029-988B-85	Sequence 85, Appl
5	31	100.0	98	US-10-032-037B-85	Sequence 85, Appl
6	31	100.0	98	US-10-032-423A-85	Sequence 85, Appl
7	31	100.0	98	US-10-194-975-15	Sequence 15, Appl
8	31	100.0	117	US-09-982-992A-6	Sequence 6, Appl
9	31	100.0	120	US-09-995-523-8	Sequence 8, Appl
10	31	100.0	241	US-09-880-748-1889	Sequence 1889, Ap
11	31	100.0	247	US-09-880-748-1470	Sequence 1470, Ap
12	31	100.0	247	US-10-322-673-48	Sequence 48, Appl
13	31	100.0	251	US-09-880-748-1594	Sequence 1594, Ap
14	31	100.0	312	US-10-288-917-10	Sequence 10, Appl
15	31	100.0	312	US-10-052-798-10	Sequence 10, Appl

16	31	100.0	395	10	US-09-738-626-4698	Sequence 4698, Ap
17	28	90.3	5	9	US-09-253-794-23	Sequence 23, Appl
18	28	90.3	5	11	US-09-977-797A-76	Sequence 76, Appl
19	28	90.3	20	12	US-10-032-221B-30	Sequence 30, Appl
20	28	90.3	46	9	US-09-864-761-48095	Sequence 48095, A
21	28	90.3	79	12	US-10-032-221B-26	Sequence 26, Appl
22	28	90.3	88	12	US-10-032-221B-33	Sequence 33, Appl
23	28	90.3	88	12	US-10-032-221B-34	Sequence 34, Appl
24	28	90.3	100	10	US-09-840-459-36	Sequence 36, Appl
25	28	90.3	118	12	US-10-270-073-3	Sequence 3, Appl
26	28	90.3	119	9	US-09-253-794-2	Sequence 2, Appl
27	28	90.3	119	9	US-09-253-794-8	Sequence 8, Appl
28	28	90.3	119	9	US-09-253-794-9	Sequence 9, Appl
29	28	90.3	119	9	US-09-253-794-10	Sequence 10, Appl
30	28	90.3	119	9	US-09-253-794-11	Sequence 11, Appl
31	28	90.3	119	9	US-09-253-794-12	Sequence 12, Appl
32	28	90.3	119	9	US-09-253-794-13	Sequence 13, Appl
33	28	90.3	119	9	US-09-253-794-14	Sequence 14, Appl
34	28	90.3	119	9	US-09-253-794-15	Sequence 15, Appl
35	28	90.3	119	9	US-09-253-794-17	Sequence 17, Appl
36	28	90.3	119	9	US-09-253-794-57	Sequence 57, Appl
37	28	90.3	119	9	US-09-253-794-58	Sequence 58, Appl
38	28	90.3	120	10	US-09-144-886-55	Sequence 55, Appl
39	28	90.3	121	10	US-09-144-886-54	Sequence 54, Appl
40	28	90.3	121	12	US-10-308-817-187	Sequence 187, Ap
41	28	90.3	124	12	US-10-032-221B-20	Sequence 20, Appl
42	28	90.3	124	12	US-10-010-729-49	Sequence 49, Appl
43	28	90.3	125	10	US-09-144-886-51	Sequence 51, Appl
44	28	90.3	125	10	US-09-144-886-52	Sequence 52, Appl
45	28	90.3	125	10	US-09-144-886-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-09-832-312-61
; Sequence 61, Application US/09832312
; Patent No. US20010049829A1
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-61
Query Match 100.0%; Score 31; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 SYMS 5
Db 1 SYMS 5
RESULT 2
US-09-829-495-61
; Sequence 61, Application US/09829495

```
Publication No. US20040001826A1
GENERAL INFORMATION:
APPLICANT: Busfield SJ
APPLICANT: Valleval J
APPLICANT: Tandrot-Petrus M
APPLICANT: Valinckenker W
APPLICANT: Gail DS
APPLICANT: Olan MD
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-234
CURRENT APPLICATION NUMBER: US/09/829,495
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 09/610,118
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/503,387
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 09/454,824
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 09/345,468
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 61
LENGTH: 5
TYPE: PRT
ORGANISM: Homo sapiens
US-09-829-495-61
```

```
Query Match          100.0%; Score 31; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SYMMS 5
        |||||
Db       1 SYMMS 5
```

```
RESULT 3
US-10-308-817-55.
Sequence 55, Application US/10308817
Publication No. US20030219861A1
GENERAL INFORMATION:
APPLICANT: Rother, Russell
APPLICANT: Mu, Dayang
TITLE OF INVENTION: HYBRID ANTIBODIES
FILE REFERENCE: 1087-37
CURRENT APPLICATION NUMBER: US/10/308,817
CURRENT FILING DATE: 2002-12-03
NUMBER OF SEQ ID NOS: 195
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55
LENGTH: 98
TYPE: PRT
ORGANISM: human
US-10-308-817-55
```

```
Query Match          100.0%; Score 31; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SYMMS 5
        |||||
Db       31 SYMMS 35
```

```
RESULT 4
US-10-029-988B-85
Sequence 85, Application US/10029988B
Publication No. US20040001839A1
GENERAL INFORMATION:
APPLICANT: Bio-Technology General Corp.
TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPTOPES CONTAINING SULFATED
TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPTOPES, AND USES THEREOF
```

```
FILE REFERENCE: 10793/46
CURRENT APPLICATION NUMBER: US/10/029,988B
CURRENT FILING DATE: 2001-12-31
PRIOR APPLICATION NUMBER: 60/258,948
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 204
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 85
LENGTH: 98
TYPE: PRT
ORGANISM: Homo sapiens
US-10-029-988B-85
```

```
Query Match          100.0%; Score 31; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SYMMS 5
        |||||
Db       31 SYMMS 35
```

```
RESULT 5
US-10-032-037B-85
Sequence 85, Application US/10032037B
Publication No. US20040001822A1
GENERAL INFORMATION:
APPLICANT: Bio-Technology General Corp.
TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPTOPES CONTAINING SULFATED
TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPTOPES, AND USES THEREOF
FILE REFERENCE: 10793/44
CURRENT APPLICATION NUMBER: US/10/032,037B
CURRENT FILING DATE: 2001-12-31
PRIOR APPLICATION NUMBER: 60/258,948
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 204
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 85
LENGTH: 98
TYPE: PRT
ORGANISM: Homo sapiens
US-10-032-037B-85
```

```
Query Match          100.0%; Score 31; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SYMMS 5
        |||||
Db       31 SYMMS 35
```

```
RESULT 6
US-10-032-423A-85
Sequence 85, Application US/10032423A
Publication No. US20040002450A1
GENERAL INFORMATION:
APPLICANT: Bio-Technology General Corp.
TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPTOPES CONTAINING SULFATED
TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPTOPES, AND USES THEREOF
FILE REFERENCE: 10793/45
CURRENT APPLICATION NUMBER: US/10/032,423A
CURRENT FILING DATE: 2001-12-31
PRIOR APPLICATION NUMBER: 60/258,948
PRIOR FILING DATE: 12/28/2000
NUMBER OF SEQ ID NOS: 204
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 85
LENGTH: 98
TYPE: PRT
ORGANISM: Homo sapiens
US-10-032-423A-85
```


Query Match 100.0%; Score 31; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 7
US-10-194-975-15
; Sequence 15, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-15

Query Match 100.0%; Score 31; DB 15; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 8
US-09-982-992A-6
; Sequence 6, Application US/09982992A
; Patent No. US2002016437A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M. et al.
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE MAP PROTEIN AND METHOD OF USE IN TRE
; FILE REFERENCE: P069220S02/BAS
; CURRENT APPLICATION NUMBER: US/09/982,992A
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/277,287
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/241,832
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-982-992A-6

Query Match 100.0%; Score 31; DB 10; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 9
US-09-995-529-8
; Sequence 8, Application US/09995529

; Publication No. US2003009655A1
; GENERAL INFORMATION:
; APPLICANT: Watkins, Jeffrey D.
; APPLICANT: Huse, William D.
; APPLICANT: Tang, Ying
; TITLE OF INVENTION: Humanized Collagen Antibodies and
; FILE REFERENCE: P-IX 4976
; CURRENT APPLICATION NUMBER: US/09/995,529
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-529-8

Query Match 100.0%; Score 31; DB 11; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 10
US-09-880-748-1889
; Sequence 1889, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1889
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1889

Query Match 100.0%; Score 31; DB 11; Length 241;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 11
US-09-880-748-1470
; Sequence 1470, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748

;; CURRENT FILING DATE: 2001-06-15
;; PRIOR APPLICATION NUMBER: 60/212,210
;; PRIOR FILING DATE: 2000-06-15
;; PRIOR APPLICATION NUMBER: 60/240,816
;; PRIOR FILING DATE: 2000-10-17
;; PRIOR APPLICATION NUMBER: 60/276,248
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/277,379
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/293,499
;; PRIOR FILING DATE: 2001-05-25
;; NUMBER OF SEQ ID NOS: 3239
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 1470
;; LENGTH: 247
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-880-748-1470

Query Match 100.0%; Score 31; DB 11; Length 247;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
|||
Db 31 SYMMS 35

RESULT 12
US-10-322-673-48
;; Sequence 48; Application US/10322673
;; Publication No. US20030180296A1
;; GENERAL INFORMATION:
;; APPLICANT: Salcedo et al.
;; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to TRAIL
;; FILE REFERENCE: PF585
;; CURRENT APPLICATION NUMBER: US/10/322,673
;; PRIOR APPLICATION NUMBER: 60/341,237
;; PRIOR FILING DATE: 2001-12-20
;; PRIOR APPLICATION NUMBER: 60/369,877
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: 60/384,828
;; PRIOR FILING DATE: 2002-06-04
;; PRIOR APPLICATION NUMBER: 60/396,591
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/403,370
;; PRIOR FILING DATE: 2002-08-15
;; PRIOR APPLICATION NUMBER: 60/425,737
;; PRIOR FILING DATE: 2002-11-13
;; NUMBER OF SEQ ID NOS: 72
;; SEQ ID NO 48
;; LENGTH: 247
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: CM013F04 scFv
US-10-322-673-48

Query Match 100.0%; Score 31; DB 12; Length 247;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
|||
Db 31 SYMMS 35

RESULT 13
US-09-880-748-1594
;; Sequence 1594; Application US/09880748
;; Publication No. US20030059937A1

;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
;; FILE REFERENCE: PF523
;; CURRENT APPLICATION NUMBER: US/09/880,748
;; CURRENT FILING DATE: 2001-06-15
;; PRIOR APPLICATION NUMBER: 60/212,210
;; PRIOR FILING DATE: 2000-06-15
;; PRIOR APPLICATION NUMBER: 60/240,816
;; PRIOR FILING DATE: 2000-10-17
;; PRIOR APPLICATION NUMBER: 60/276,248
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/277,379
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/293,499
;; PRIOR FILING DATE: 2001-05-25
;; NUMBER OF SEQ ID NOS: 3239
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 1594
;; LENGTH: 251
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-880-748-1594

Query Match 100.0%; Score 31; DB 11; Length 251;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMMS 5
|||
Db 31 SYMMS 35

RESULT 14
US-10-288-917-10
;; Sequence 10; Application US/10288917
;; Publication No. US20030148455A1
;; GENERAL INFORMATION:
;; APPLICANT: Adams, Camilla W.
;; Ashkenazi, Avi J.
;; Chuntharapai, Anan
;; Kim, Kyung J.
;; TITLE OF INVENTION: Apo-2 Receptor
;; NUMBER OF SEQUENCES: 14
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 1 DNA Way
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WinPatln (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/288,917
;; FILING DATE: 06-No. US20030148455A1-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 10/052798
;; FILING DATE: 02-NOV-2001
;; APPLICATION NUMBER: 09/079029
;; FILING DATE: 14-MAY-1998
;; APPLICATION NUMBER: 60/074119
;; FILING DATE: 09-FEB-1998
;; APPLICATION NUMBER: 60/046615
;; FILING DATE: 15-MAY-1997

ATTORNEY/AGENT INFORMATION:
NAME: Marechany, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P1101R2D1C1

Search completed: January 12, 2004, 07:19:52
Job time : 15 secs

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-5416

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 312 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-288-917-10

Query Match

100.0%; Score 31; DB 12; Length 312;

Best Local Similarity 100.0%; Pred. No. 7.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5

DB 70 SYMS 74

RESULT 15

US-10-052-798-10

Sequence 10, Application US/10052798

Publication No. US20020150985A1

GENERAL INFORMATION:

APPLICANT: Adams, Camilia W.

Ashkenazi, Avi J.

Chuntharapai, Anan

Kim, Kyung J.

TITLE OF INVENTION: Apo-2 Receptor

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPacIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/052,798

FILING DATE: 02-No. US20020150985A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/079,029

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Marschang, Diane L.

REGISTRATION NUMBER: 35,600

REFERENCE/DOCKET NUMBER: P1101R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-5416

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 312 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-052-798-10

Query Match

100.0%; Score 31; DB 14; Length 312;

Best Local Similarity 100.0%; Pred. No. 7.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5

DB 70 SYMS 74

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:49 ; Search time 15.4062 Seconds
(without alignments)
106.117 Million cell updates/sec

Title: US-09-829-495-62

Perfect score: 89
Sequence: 1 NIKODSEKXYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	96.6	147	2 I37780	Ig variable region
2	82	92.1	97	2 PH0872	Ig heavy chain V r
3	82	92.1	98	2 PL0123	Ig heavy chain V r
4	82	92.1	110	2 PH1655	Ig heavy chain V r
5	82	92.1	114	2 S36280	Ig heavy chain V r
6	82	92.1	117	2 S78486	Ig heavy chain V r
7	82	92.1	117	2 S17079	Ig heavy chain V r
8	82	92.1	122	2 S31675	Ig heavy chain V r
9	82	92.1	136	2 S31587	Ig heavy chain V r
10	82	92.1	139	2 I37781	Ig variable region
11	79	88.8	140	2 S22657	Ig heavy chain pre
12	77	86.5	123	2 S31509	Ig heavy chain - h
13	67	75.3	136	2 S16847	Ig heavy chain V r
14	67	75.3	136	2 S60296	Ig heavy chain V r
15	66	74.2	98	2 S29543	Ig heavy chain var
16	66	74.2	117	2 S36259	Ig heavy chain V r
17	66	74.2	134	2 S31688	Ig heavy chain V r
18	66	74.2	135	2 S31598	Ig heavy chain V r
19	66	74.2	140	2 S70442	Ig heavy chain pre
20	65	73.0	121	2 PH1661	Ig heavy chain V r
21	64	71.9	94	2 PL0120	Ig heavy chain V r
22	64	71.9	97	2 S44115	Ig heavy chain V r
23	64	71.9	98	2 PL0116	Ig heavy chain V r
24	64	71.9	98	2 S29546	Ig heavy chain V r
25	64	71.9	109	2 PH1644	Ig heavy chain V r
26	64	71.9	109	2 PH1646	Ig heavy chain V r
27	64	71.9	111	2 PH1643	Ig heavy chain V r
28	64	71.9	111	2 PH1645	Ig heavy chain V r
29	64	71.9	113	2 S38490	Ig heavy chain - h

30	64	71.9	114	2 S46390	Ig heavy chain V r
31	64	71.9	114	2 S46392	Ig heavy chain V r
32	64	71.9	117	2 S36270	Ig heavy chain V r
33	64	71.9	118	2 S31677	Ig heavy chain V r
34	64	71.9	118	2 S31116	Ig heavy chain - h
35	64	71.9	119	2 P36005	Ig heavy chain V r
36	64	71.9	120	2 S31112	Ig heavy chain - h
37	64	71.9	121	2 S19666	Ig heavy chain V r
38	64	71.9	121	2 G36005	Ig heavy chain V r
39	64	71.9	122	2 E36005	Ig heavy chain V r
40	64	71.9	122	2 S31119	Ig heavy chain - h
41	64	71.9	123	2 S38493	Ig heavy chain - h
42	64	71.9	130	2 PL0098	Ig heavy chain pre
43	64	71.9	132	2 S31603	Ig heavy chain V r
44	64	71.9	134	2 S31679	Ig heavy chain V r
45	64	71.9	139	2 S31674	Ig heavy chain V r

ALIGNMENTS

RESULT 1

I37780
Ig variable region (VDJ) (clone T20-11) - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #ext_change 23-Jul-1999
R/Demission: C.; Chastagner, P.; There, J.; Zouali, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
A/Title: Somatic diversification in the heavy chain variable region genes expressed by
A/Reference number: A36876; MUID:94119917; PMID:8290556
A/Accession: I37780
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-147 <RES>
A/Cross-references: EMBL:X67943; NID:933578; PID:CAA48130.1; PID:933579
C/Superfamily: Immunoglobulin V region; Immunoglobulin homology
F/28-11/Domain: Immunoglobulin homology <IMM>

Query Match

Best Local Similarity 96.6%; Score 86; DB 2; Length 147;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIKODSEKXYADSVRG 17
DB 63 NIKODSEKXYADSVRG 79

RESULT 2

PH0872
Ig heavy chain V region (anti-DNA, III-3R) - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #ext_change 16-Aug-1996
A/Accession: PH0872
J/Rambhimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A/Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idioctyp
A/Reference number: PH0862; MUID:92078875; PMID:1660528
A/Accession: PH0872
A/Molecule type: DNA
A/Residues: 1-97 <MAN>
A/Note: a "G" was inserted at the position of between 72nd and 73rd of the DNA sequence
C/Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that bear
C/Superfamily: Immunoglobulin V region; Immunoglobulin homology
C/Keywords: heterotetramer; Immunoglobulin
F/14-97/Domain: immunoglobulin homology <IMM>
F/30-35/Region: complementarity-determining 1
F/49-66/Region: complementarity-determining 2

Query Match

Best Local Similarity 92.1%; Score 82; DB 2; Length 97;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKYYADSVRG 17
|||||
Db 49 NIKODSEKYYVDSVKG 65

RESULT 3

PL0123
Ig heavy chain V-III region (TD-Vr) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999

C/Accession: PL0123; S26897

R/Bird: U.; Gallili, N.; Link, M.; Stites, D.; Sklar, J.

J. Exp. Med. 168, 229-245, 1988

A/Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin

C/Keywords: heterotetramer; immunoglobulin

A/Reference number: S26885; PMID:9302117; PMID:1404388

A/Accession: PL0123

A/Molecule type: mRNA

A/Residues: 1-98 <BIR>

A/Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL

A/Note: the sequence shows the V region (TD-Vr) from a nonproductive DNA rearrangement

R/Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.

J. Mol. Biol. 227, 776-798, 1992

A/Title: The repertoire of human germline V(H) sequences reveals about fifty groups of

A/Reference number: S26885; PMID:9302117; PMID:1404388

A/Accession: S26897

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-98 <TOM>

A/Cross-references: EMBL:Z12354; NID:932930; PIDN:CA78224.1; PID:932931

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin

F/15-98/Domain: immunoglobulin homology <IMM>

F/15-98/Region: complementarity-determining 1

F/49-65/Region: complementarity-determining 2

Query Match

Best Local Similarity 92.1%; Score 82; DB 2; Length 98;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKYYADSVRG 17
|||||
Db 50 NIKODSEKYYVDSVKG 66

RESULT 4

PH1655
Ig heavy chain V region (clone 2B8) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996

C/Accession: PH1655

R/Hillson, J.L.; Karr, N.S.; Opliger, I.R.; Mannik, M.; Sasso, E.H.

J. Exp. Med. 178, 331-336, 1993

A/Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphylo

C/Reference number: PH1642; PMID:9301610; PMID:8315388

A/Accession: PH1655

A/Molecule type: mRNA

A/Residues: 1-110 <HTL>

A/Experimental source: B cell

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F/7-90/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 92.1%; Score 82; DB 2; Length 110;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKYYADSVRG 17
|||||
Db 42 NIKODSEKYYVDSVKG 58

RESULT 5

S36280

Ig heavy chain V region (clone alpha-FOG1-A3) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999

C/Accession: S36280

R/Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J

EMBO J. 12, 725-734, 1993

A/Title: Human anti-self antibodies with high specificity from phage display libraries.

A/Reference number: S36256; PMID:93178448; PMID:7679990

A/Accession: S36280

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-114 <GRI>

A/Cross-references: EMBL:Z18822

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F/15-98/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 92.1%; Score 82; DB 2; Length 114;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKYYADSVRG 17
|||||
Db 50 NIKODSEKYYVDSVKG 66

RESULT 6

S78486
Ig heavy chain V region (clone FL13-28) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 21-Jan-2000

C/Accession: S78486; S31115

R/Raaphorst, F.M.

submitted to the EMBL Data Library, October 1991

A/Reference number: S78486

A/Accession: S78486

A/Molecule type: mRNA

A/Residues: 1-117 <RAA>

A/Cross-references: EMBL:X62965

A/Experimental source: clone FL13-28

R/Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurmar

Eur. J. Immunol. 22, 247-251, 1992

A/Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comple

C/Reference number: S31104; PMID:92111633; PMID:1730252

A/Accession: S31115

A/Molecule type: mRNA

A/Residues: 1-4, 'L', 6-32, 'G', 34-52, 'E', 54-73, 'K', 75-97 <RAW>

A/Cross-references: EMBL:X62965

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F/15-98/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 92.1%; Score 82; DB 2; Length 117;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKYYADSVRG 17
|||||
Db 50 NIKODSEKYYVDSVKG 66

RESULT 7

S17079

Ig heavy chain V-gene (clone HHG19) - human

C/Species: Homo sapiens (man)

C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999

C/Accession: S60299; S17079

R/Kuepers, R.; Fischer, U.; Rajewsky, K.; Gause, A.

Immunol. Lett. 34, 57-62, 1992

A/Title: Immunoglobulin heavy and light chain gene sequences of a human CD5 positive im

A/Reference number: S60295; PMID:93122853; PMID:1282458

A/Accession: S60299

A/Status: preliminary

C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:23-106/Domain: immunoglobulin homology <IMM>

Query Match 86.5%; Score 77; DB 2; Length 123;
Best Local Similarity 82.4%; Pred. No. 1.3e-05;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NIKDGSERYADSVRG 17
:|||||:
Db 58 NIKDGSERYADSVRG 74

RESULT 13

IG heavy chain V region - human
C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Aug-1996
C/Accession: S16847
R:Knepfer, R.
submitted to the EMBL Data Library, September 1991
A/Reference number: S16847
A/Accession: S16847
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-136 <KUE>
A/Cross-references: EMBL:X62126
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 75.3%; Score 67; DB 2; Length 136;
Best Local Similarity 68.8%; Pred. No. 0.00072;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKDGSERYADSVRG 17
:|||||:
Db 70 VKDGSERYADSVRG 85

RESULT 14

S60296
IG heavy-chain variable region (clone WHR) precursor - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Jul-1996 #sequence_revision 27-Feb-1997 #text_change 21-Jan-2000
C/Accession: S60296
R:Knepfer, R.; Fischer, U.; Rajewsky, K.; Gause, A.
Immunol. Lett. 34, 57-62, 1992
A/Title: Immunoglobulin heavy and light chain gene sequences of a human CD5 positive imm
A/Reference number: S60295; MUID:9312853; PMID:1282498
A/Accession: S60296
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-136 <KUE>
A/Cross-references: EMBL:X62126; NID:938336; PIDN:CAA44057.1; PID:938337
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 75.3%; Score 67; DB 2; Length 136;
Best Local Similarity 68.8%; Pred. No. 0.00072;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKDGSERYADSVRG 17
:|||||:
Db 70 VKDGSERYADSVRG 85

RESULT 15

S29543
IG heavy chain V region (COS 3) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jan-1994 #sequence_revision 17-Nov-1995 #text_change 23-Jul-1999
C/Accession: S29543

R:Tomlinson, M.; Walter, G.; Cook, G.P.; Winter, G.
submitted to the EMBL Data Library, October 1992
A/Reference number: S29543
A/Accession: S29543
A/Molecule type: DNA
A/Residues: 1-98 <TCM>

A/Cross-references: EMBL:217389; NID:932835; PIDN:CAA78994.1; PID:932836
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 74.2%; Score 66; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.00074;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKDGSERYADSVRG 17
:|||||:
Db 51 IKDGSERYADSVRG 66

Search completed: January 12, 2004, 07:02:57
Job time : 15.4062 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:49 ; Search time 9.29688 Seconds
(without alignments)
85.992 Million cell updates/sec

Title: US-09-829-495-62

Perfect score: 89

Sequence: 1 NIKODGSEKRYADSVRG 17

Scoring table: BLOSUM62

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	70.8	116	1 HV3T_HUMAN	P01781 homo sapien
2	57	64.0	122	1 HV3G_HUMAN	P01768 homo sapien
3	56	62.9	118	1 HV3V_HUMAN	P80419 homo sapien
4	55	61.8	119	1 HV3L_HUMAN	P01772 homo sapien
5	55	61.8	126	1 HV3K_HUMAN	P01772 homo sapien
6	50	56.2	119	1 HV3N_HUMAN	P01774 homo sapien
7	50	56.2	119	1 HV3N_HUMAN	P01775 homo sapien
8	49	55.1	119	1 HV3I_HUMAN	P01770 homo sapien
9	47	52.8	122	1 HV3A_HUMAN	P01762 homo sapien
10	47	52.8	122	1 HV3H_HUMAN	P01763 homo sapien
11	46	51.7	114	1 HV3B_HUMAN	P19180 carassius a
12	46	51.7	117	1 HV03_CARAU	P01771 homo sapien
13	46	51.7	121	1 HV3U_HUMAN	P01771 homo sapien
14	45	50.6	98	1 HV57_MOUSE	P18528 mus musculu
15	45	50.6	116	1 HV05_CARAU	P19181 carassius a
16	44	49.4	115	1 HV3F_HUMAN	P01767 homo sapien
17	44	49.4	121	1 YE04_MERJA	O58799 methanococ
18	43	48.3	117	1 HV3C_HUMAN	P01764 homo sapien
19	43	48.3	120	1 HV3E_HUMAN	P01766 homo sapien
20	42	47.2	117	1 HV53_MOUSE	P18524 mus musculu
21	42	47.2	258	1 YC09_YEAST	P25637 saccharomyc
22	41	46.1	117	1 HV01_CAICR	P01813 calman croc
23	41	46.1	357	1 Y046_MYCTU	P71703 mycobacteri
24	41	46.1	623	1 PBP2_SALTY	P74872 salmonella
25	41	46.1	877	1 DYN_DROME	P27619 drosophila
26	41	46.1	1085	1 S124_HUMAN	O91695 homo sapien
27	41	46.1	1085	1 S124_MOUSE	O91688 mus musculu
28	41	46.1	1085	1 S124_RABIT	O28677 oryctolagus
29	41	46.1	1085	1 S124_RAT	O63632 rattus norv
30	41	46.1	1150	1 S126_HUMAN	O91695 homo sapien
31	41	46.1	1150	1 S126_MOUSE	O92444 mus musculu
32	41	46.1	4447	1 PKXK_BACSU	P40803 bacillus su
33	40.5	45.5	197	1 BINL_STAAU	P18358 staphylococ

34	40	44.9	97	1 HV56_MOUSE	P18527 mus musculu
35	40	44.9	117	1 HV54_MOUSE	P18525 mus musculu
36	40	44.9	117	1 HV55_MOUSE	P18526 mus musculu
37	40	44.9	198	1 JDP1_MOUSE	O91022 mus musculu
38	40	44.9	229	1 GPMA_LISTIN	O92958 listeria in
39	40	44.9	229	1 GPMA_LISTMO	O85571 listeria mo
40	40	44.9	231	1 PSBQ_ONOVI	O22591 onobrychis
41	40	44.9	318	1 FMT_HAEIN	P44787 haemophilus
42	40	44.9	419	1 RECA_MYCS	O91414 mycobacteri
43	40	44.9	463	1 YCE8_YEAST	P25380 saccharomyc
44	39	43.8	113	1 HV31_MOUSE	P01800 mus musculu
45	39	43.8	117	1 HV58_MOUSE	P18529 mus musculu

ALIGNMENTS

RESULT 1
HV3T_HUMAN STANDARD; PRT; 116 AA.
AC P01781;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 19 heavy chain V-II region GAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RX MEDLINE=75059123; PubMed=4803843;
RA Watanabe S., Barnikol H.U., Horn J., Hilschmann N.;
RT "The primary structure of a monoclonal IgM-immunoglobulin
(macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-
type), subgroup H II. Architecture of the complete IgM-molecule.";
RT Hoppe-Sejler's Z. Physiol. Chem. 354:1505-1509(1973).
RN (2)
RP REVISION TO 28-33.
RA Hilschmann N.,
RL Submitted (JUN-1975) to the PIR data bank.
CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S
MACROGLOBULIN.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR: A02064; M3HUGL.
DR HSSP: P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 116 112 IG-LIKE.
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9AAAAA1282 CRC64;
Query Match 70.8%; Score 63; DB 1; Length 116;
Best local Similarity 58.8%; Pred. No. 0.00088;
Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
OY 1 NIKODGSEKRYADSVRG 17
DB 50 NIKZBGSZZBYDVSXG 66
RESULT 2
HV3G_HUMAN STANDARD; PRT; 122 AA.
ID HV3G_HUMAN
AC P01768;

DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region CAM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=61013859; PubMed=6774332;
 RA Lehman D.W., Putnam F.W.;
 RT "Amino acid sequence of the variable region of a human mu chain:
 location of a possible JH segment."
 RT Proc. Natl. Acad. Sci. U.S.A. 77:3239-3243(1980).
 CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM THE PLASMA OF A
 CC PATIENT WITH MACROGLOBULINEMIA.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR HSP; P01772; 2F84.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003823; F:antigen binding activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Immunoglobulin V region; Pyroliidone carboxylic acid.
 FT DOMAIN 1 112
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT NON_TER 122 122
 FT SEQUENCE 122 AA; 13668 MW; A42D0F17D252FIC2 CRC64;
 SQ
 Query Match 64.0%; Score 57; DB 1; Length 122;
 Best Local Similarity 56.2%; Pred. No. 0.0099;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 2 IKODGSEKTYADSVRG 17
 DB 51 ISYGBBKYYABSVKG 66
 RESULT 3
 ID HV3L_HUMAN STANDARD; PRT; 118 AA.
 AC P80419;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region GAR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=95255299; PubMed=7737190;
 RA Scoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
 RT "Characterization of the two unique human anti-flavin monoclonal
 immunoglobulins."
 RT Immunoglobulins.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR HSP; P01772; 2F84.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003823; F:antigen binding activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 110
 FT MOD RES 1 1 IG-LIKE.
 FT NON_TER 118 118
 FT SEQUENCE 118 AA; 13087 MW; 6C21D810ED1B6D1F CRC64;
 SQ
 Query Match 62.9%; Score 56; DB 1; Length 118;
 Best Local Similarity 62.5%; Pred. No. 0.014;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NIKODGSEKTYADSVR 16
 DB 50 NIKPDETEKYSYSVR 65
 RESULT 4
 ID HV3L_HUMAN STANDARD; PRT; 119 AA.
 AC P01773;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region BUR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE (MYELOMA PROTEIN BUR).
 RX MEDLINE=79151016; PubMed=107164;
 RA Putnam F.W., Liu Y.-S.V., Low T.L.K.;
 RT "Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal
 RT IgA1 protease, digestion, Fab and Fc fragments, and the complete
 RT amino acid sequence of the alpha 1 heavy chain."
 RT J. Biol. Chem. 254:2865-2874(1979).
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR HSP; P01772; 2F84.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003823; F:antigen binding activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Immunoglobulin V region; Glycoprotein; Pyroliidone carboxylic acid.
 FT DOMAIN 1 112
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 22 96
 FT CARBOHYD 28 28
 FT NON_TER 119 119
 FT SEQUENCE 119 AA; 12981 MW; 12A709A75344D024 CRC64;
 SQ
 Query Match 61.8%; Score 55; DB 1; Length 119;
 Best Local Similarity 68.8%; Pred. No. 0.021;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 IKODGSEKTYADSVRG 17
 DB 51 ISYGBBKYYADSVRG 66
 RESULT 5
 ID HV3K_HUMAN STANDARD; PRT; 126 AA.
 AC P01772;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region KOL.

OC	Homo sapiens (Human).
CC	Euryarchaeota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX	NCL_taxonomy=9606;
RN	[1]
RP	SEQUENCE, AND DISULFIDE BONDS.
RX	MEDLINE=63289131; PubMed=6884994;
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT	"Three-dimensional structural structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.",
RL	Hoepf-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN	[2]
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX	MEDLINE=61072295; PubMed=7441755;
RA	Marquart M., Deisenhofer J., Huber R., Palm W.;
RT	"Crystallographic refinement and atomic models of the intact immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A
RL	J. Mol. Biol. 141:369-391(1980).
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR	PDB; A02055; GIMDL.
DR	PDB; 2FB4; 12-JUL-89.
DR	PDB; 2IG2; 12-JUL-89.
DR	GO; GO:0005576; C:extracellular; NAS.
DR	GO; GO:0003823; F:antigen binding activity; NAS.
DR	GO; GO:0006955; P:immune response; NAS.
DR	InterPro; IPRO07110; IG-like.
DR	InterPro; IPRO03066; IG_MHC.
DR	InterPro; IPRO03596; IG_v.
DR	Pfam; PF00047; IG; 1.
DR	SMART; SMO0406; IGv; 1.
DR	PROSITE; PS50835; IG LIKE; 1.
KM	Immunoglobulin V region; 3D-structure; Pyrrolidone carboxylic acid.
FT	DOMAIN 1 112 IG-LIKE.
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT	DISULFID 22 96
FT	DISULFID 105 110
FT	STRAND 3 7
FT	STRAND 11 12
FT	TURN 14 15
FT	STRAND 18 25
FT	HELIX 29 31
FT	STRAND 34 39
FT	TURN 41 42
FT	STRAND 45 51
FT	TORN 53 54
FT	STRAND 58 60
FT	HELIX 62 64
FT	STRAND 65 65
FT	TURN 66 67
FT	STRAND 68 73
FT	TORN 74 77
FT	STRAND 78 83
FT	HELIX 88 90
FT	STRAND 92 99
FT	TORN 106 106
FT	STRAND 107 108
FT	STRAND 109 109
FT	STRAND 113 116
FT	STRAND 120 124
FT	NON_TER 126 126
QO	SEQUENCE 126 AA; 13718 MW; E4D71B52B1EF8776 CRC64;

Query Match	61.8%	Score 55;	DB 1;	Length 126;
Best Local Similarity	62.5%	Pred. No. 0.02;		
Matches	10;	Conservative	4;	Mismatches 2;
				Indels 0;
				Gaps 0
QY	2	IKODGSEKRYADSVRG	17	
Db	51	IMDDGSDQHYADSVKG	66	
RESULT	6			

ID	HW3M_HUMAN	STANDARD;	PRT;	119 AA.
AC	HW3M_HUMAN	STANDARD;		
DT	21-JUL-1986 (rel. 01, Created)			
DT	21-JUL-1986 (rel. 01, Last sequence update)			
DT	15-SEP-2003 (rel. 42, Last annotation update)			
DE	IG heavy chain V-II region POM.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=75046755; PubMed=4139708;			
RA	Capra J.D., Kehoe J.M.;			
RT	"Structure of antibodies with shared idiotypy: the complete sequence			
RT	of the heavy chain variable regions of two immunoglobulin M			
RT	anti-gamma globulins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).			
CC	-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA			
CC	GLOBULIN ACTIVITY.			
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like domain.			
DR	PIR; A02057; M3HUPM.			
DR	HSSP; P01772; 2P84.			
DR	GO: GO:0005576; C:extracellular; NAS.			
DR	GO: GO:0003823; P:antigen binding activity; NAS.			
DR	GO: GO:0006955; P:immune response; NAS.			
DR	InterPro; IPR007110; IG-1-like.			
DR	InterPro; IPR003006; IG_MHC.			
DR	InterPro; IPR003596; IG_V.			
DR	Pfam; PF00047; IG_1.			
DR	SMART; SM00406; IGV; 1.			
DR	PROSITE; PS50835; IG_LINE; 1.			
DR	Immunoglobulin V region.			
FT	DOMAIN	1	112	
FT	FT	54	54	
FT	VARIANT			IG-LIKE.
FT				N -> D (PROBABLY DUE TO DEAMIDATION
FT				DURING ISOLATION).
FT				/FTId=VAR_003966.
SO	SEQUENCE	119 AA;	119 MM;	2E018AF4DCBB2610 CRC64;
DR	NON TER	119	119	
DR	Best Match	56.2%;	Score 50;	DB 1; Length 119;
DR	Best Local Similarity	57.1%;	Pred. No. 0.15;	
DR	Matches	8; Conservative	5; Mismatches	1; Indels 0; Gaps
QY	4 ODGSEKYADSVRG 17			
Db	53 ENGNDKHYADSYNG 66			
RESULT 7				
HW3N_HUMAN				
ID	HW3N_HUMAN	STANDARD;	PRT;	119 AA.
AC	HW3N_HUMAN	STANDARD;		
DT	21-JUL-1986 (rel. 01, Created)			
DT	21-JUL-1986 (rel. 01, Last sequence update)			
DT	15-SEP-2003 (rel. 42, Last annotation update)			
DE	IG heavy chain V-III region LAY.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=75046755; PubMed=4139708;			
RA	Capra J.D., Kehoe J.M.;			
RT	"Structure of antibodies with shared idiotypy: the complete sequence			
RT	of the heavy chain variable regions of two immunoglobulin M			
RT	anti-gamma globulins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).			
CC	-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA			
CC	GLOBULIN ACTIVITY.			
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like domain.			

```

SQ SEQUENCE 119 AA; 12953 MW; 2E018AF4DCEB2610 CRC64;

Query Match 56.2%; Score 50; DB 1; Length 119;
Best Local Similarity 57.1%; Pred. No. 0.15;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0.

QY 4 ODGEKRYADSVRG 17
   ::|::|::|::|::|
Db 53 ENGNDKRYADSVNG 66

RESULT 7
HV3N_HUMAN
ID HV3N_HUMAN STANDARD; PRT; 119 AA.
AC P01775;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DS Ig heavy chain V-II region IAY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RZ SEQUENCE.
RX MEDLINE=75046755; PubMed=4139708;
RA Capra J.D., Kehoe J.M.;
RT "Structure of antibodies with shared idiotypy: the complete sequence
RT of the heavy chain variable regions of two immunoglobulin M
RL anti-gamma globulins.";
RL Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).
-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.

```

DR PIR; A02058; M3HULY.
 DR HSSP; P01772; 2FB4.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; F:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_MHC.
 DR Pfam; PF00047; Ig_v.
 DR SMART; SM00406; Ig; 1.
 DR PROSITE; PS50835; IG-LIKE; 1.
 DR Immunoglobulin V region.
 FT DOMAIN 1 112 IG-LIKE.
 FT NON TER 119 119
 SQ SEQUENCE 119 AA; 12858 MW; D6338098794DCFE CRC64;
 Query Match 56.2%; Score 50; DB 1; Length 119;
 Best Local Similarity 57.1%; Pred. No. 0.15;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 4 DQSEKYYADSVRG 17
 Db 53 ENGNDKHYADSVNG 66

RESULT 8
 HV31 HUMAN
 ID HV31 HUMAN STANDARD; PRT; 119 AA.
 AC P01770;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region NIE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=77070269; PubMed=826475;
 RA Ponstingl H., Hilschmann N.;
 RT "The rule of antibody structure. The primary structure of a
 RT monoclonal IgG1 immunoglobulin (myeloma protein Nle), III. The
 RT chymotryptic peptides of the H-chain, alignment of the tryptic
 RT peptides and discussion of the complete structure.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IG1 MYELOMA
 CC PROTEIN.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC PIR; A91668; GIHUNI.
 DR HSSP; P01772; 2FB4.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; Ig; 1.
 DR PROSITE; PS50835; IG-LIKE; 1.
 DR Immunoglobulin V region; Pyroliidone carboxylic acid.
 FT DOMAIN 1 112 IG-LIKE.
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD RES 1 1
 FT DISULFID 22 96

FT NON TER 119 119
 SQ SEQUENCE 119 AA; 13242 MW; C96935A6E55E165B CRC64;
 Query Match 55.1%; Score 49; DB 1; Length 119;
 Best Local Similarity 61.5%; Pred. No. 0.22;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 DQSEKYYADSVRG 17
 Db 54 BGBBKHYADSVNG 66

RESULT 9
 HV31 HUMAN
 ID HV31 HUMAN STANDARD; PRT; 122 AA.
 AC P01762;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region TRO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE (MYELOMA PROTEIN TRO).
 RX MEDLINE=76023781; PubMed=809331;
 RA Kratzin H., Altevogt P., Ruban E., Kortt A., Starosck K.,
 RA Hilschmann N.;
 RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro),
 RT II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
 RT structure of the complete IgA-molecule.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
 CC -1- MISCELLANEOUS: THE SEQUENCE OF THE C REGION IS ALSO GIVEN.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR; A02045; AHUTR.
 DR HSSP; P01772; 2FB4.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; F:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; Ig; 1.
 DR PROSITE; PS50835; IG-LIKE; 1.
 DR Immunoglobulin V region; Pyroliidone carboxylic acid.
 FT DOMAIN 1 108 IG-LIKE.
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD RES 1 1
 FT NON TER 122 122
 SQ SEQUENCE 122 AA; 13472 MW; 2E21A11DA04D80F9 CRC64;
 Query Match 52.8%; Score 47; DB 1; Length 122;
 Best Local Similarity 75.0%; Pred. No. 0.5;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 GSEKYYADSVRG 17
 Db 55 GSTLYYADSVNG 66

RESULT 10
 HV31 HUMAN
 ID HV31 HUMAN STANDARD; PRT; 122 AA.
 AC P01769;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region GA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_Taxid=9606;

```

RN [1]
RP SEQUENCE.
RX MEDLINE=74175307; PubMed=4208843;
RT Florent G., Lehman D., Putnam F.W.;
RL "The switch point in mu heavy chains of human IGM immunoglobulins.";
CC Biochemistry 13:2482-2498(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
MACROGLOBULIN.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02052; M3HUGA.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; Ig_LIKE; 1.
KW Immunoglobulin V region; Pyrrolidone carboxylic acid.
FT MOD_RES 1 112
FT NON_TER 1 122
SQ SEQUENCE 122 AA; 13166 MW; 74E5B6959EB4100A CRC64;

Query Match
Best Local Similarity 52.8%; Score 47; DB 1; Length 122;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 IKDSEKXYADSVRG 17
DB 51 ISYBGBZYAAASVKG 66

RESULT 11
H3B3_HUMAN
ID H3B3_HUMAN STANDARD; PRT; 114 AA.
AC P01763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-IIi region WEA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_Taxid=9606;
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=83273707; PubMed=6410398;
RT Goni F., Frangione B.;
RT "Amino acid sequence of the Fv region of a human monoclonal IGM
(galactose in Klebsiella polysaccharides K30 and K33).";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY
AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH
WALDENSTROM'S MACROGLOBULINEMIA.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02046; M3HUGA.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; Ig_LIKE; 1.
KW Immunoglobulin V region; Pyrrolidone carboxylic acid.
FT MOD_RES 1 112
FT NON_TER 1 122
SQ SEQUENCE 117 AA; 13220 MW; 51B625003FAECB CRC64;

```

```

FT NON_TER 114 114
SQ SEQUENCE 114 AA; 12256 MW; D88294FB418A07B7 CRC64;

Query Match
Best Local Similarity 51.7%; Score 46; DB 1; Length 114;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 6 GSEKXYADSVRG 17
DB 55 GSTIYYADSVKG 66

RESULT 12
H303_CARAU
ID H303_CARAU STANDARD; PRT; 117 AA.
AC P19180;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 3 precursor.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OC NCBI_Taxid=7957;
OX NCBI_Taxid=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88144476; PubMed=3125551;
RT Wilson M.R., Middleton D., Watt G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure
and family relationships of two genes and a pseudogene in a teleost
fish.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
CC DR EMBL; J03616; AAA50807.1; -.
CC DR PIR; A28966; A28966.
CC DR HSSP; P01772; 2FB4.
CC DR InterPro; IPR007110; Ig-like.
CC DR InterPro; IPR003006; Ig_MHC.
CC DR InterPro; IPR003596; Ig_V.
CC DR Pfam; PF00047; Ig_1.
CC DR SMART; SM00406; IGV; 1.
CC DR PROSITE; PS50835; Ig_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT CHAIN 1 19
FT SIGNAL 1 19
FT DOMAIN 20 117 IG HEAVY CHAIN V REGION 3.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 FRAMEWORK-3.
FT DOMAIN 86 117 COMPLEMENTARITY-DETERMINING-2.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117
FT MOD_RES 117
SQ SEQUENCE 117 AA; 13220 MW; 51B625003FAECB CRC64;

Query Match
Best Local Similarity 51.7%; Score 46; DB 1; Length 117;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 IKDSEKXYADSVRG 17
DB 70 IYDSDSVYADTVKG 85

RESULT 13

```

```

HV3J HUMAN
ID HV3J HUMAN STANDARD; PRT; 121 AA.
AC P01771;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-II region HIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=79124635; Pubmed=420800;
RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;
RT "Amino acid sequence of the VH region of human myeloma
RL cryoimmunoglobulin IgG H1.1."
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR HSP; A02054; G1HHL.
DR HSP; P01772; 2PB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR003006; Ig-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR POSITE; PS50835; IG_LIKE; 1.
DR Immunoglobulin V region; Pyroglutamate carboxylic acid.
FT DOMAIN 1 112 IG-LIKE.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13566 MW; 480FC53610F5DAB CRC64;

Query Match 51.7%; Score 46; DB 1; Length 121;
Best Local Similarity 61.5%; Pred. No. 0.74;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 DGESEKYADSVRG 17
Db 54 NGSRTYGGDSYVG 66

RESULT 14
HV57 MOUSE
ID HV57 MOUSE STANDARD; PRT; 98 AA.
AC P18528;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region 6.96.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/CJ;
RX MEDLINE=89279149; Pubmed=2499654;
RA Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
RT "Early onset of somatic mutation in immunoglobulin VH genes during
RT the primary immune response."
RL J. Exp. Med. 169:2007-2019(1989).
CC -1- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR HSP; J0501; HVMS96.
DR HSP; P01772; 2PB4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.

```

```

DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR POSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 288 IG-LIKE.
FT NON_TER 98 98
SQ SEQUENCE 98 AA; 11007 MW; B8644F792F8P95B CRC64;

Query Match 50.6%; Score 45; DB 1; Length 98;
Best Local Similarity 56.2%; Pred. No. 0.87;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IKDGESEKYADSVRG 17
Db 51 ISDGSYTYPPDSYVG 66

RESULT 15
HV05 CARAU
ID HV05 CARAU STANDARD; PRT; 116 AA.
AC P19181;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 5A precursor.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8814476; Pubmed=312551;
RA Wilson M.R., Middleton D., Marr G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure
RT and family relationships of two genes and a pseudogene in a teleost
RT fish."
RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
DR HSP; P01772; 2PB4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR POSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 116 IG HEAVY CHAIN V REGION 5A.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 85 116 FRAMEWORK-3.
FT DISUPID 41 114 BY SIMILARITY.
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;

Query Match 50.6%; Score 45; DB 1; Length 116;
Best Local Similarity 66.7%; Pred. No. 1;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GSESEKYADSVRG 17
Db 73 GGSYTYADSVYG 84

Search completed: January 12, 2004, 06:59:36
Job time : 10.2969 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:04 ; Search time 34.5312 Seconds
(without alignments)
127.041 Million cell updates/sec

Title: US-09-829-495-62

Perfect score: 89
Sequence: 1 NIKQDSEKXYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	92.1	95	4 Q9ULB6	Q9ULB6 homo sapien
2	66	74.2	113	4 Q9UL90	Q9UL90 homo sapien
3	64	71.9	116	4 Q9UL93	Q9UL93 homo sapien
4	64	71.9	613	4 Q8WUK1	Q8WUK1 homo sapien
5	61	68.5	122	4 Q9UL84	Q9UL84 homo sapien
6	53	59.6	121	4 Q9UL71	Q9UL71 homo sapien
7	52	58.4	147	4 Q9Y509	Q9Y509 homo sapien
8	50	56.2	71	7 Q9GJ71	Q9GJ71 salmo trutt
9	48	53.9	493	4 Q8NCL6	Q8NCL6 homo sapien
10	48	53.9	118	4 Q9UL72	Q9UL72 homo sapien
11	48	53.9	593	16 Q9ZHB9	Q9ZHB9 rickettsia
12	48	53.9	594	16 Q9ZCY6	Q9ZCY6 rickettsia
13	47.5	53.4	247	5 Q9VZF9	Q9VZF9 drosophila
14	47	52.8	342	2 Q9FSK2	Q9FSK2 rhizobium l
15	47	52.8	360	16 Q987Q4	Q987Q4 rhizobium l
16	47	52.8	597	4 Q96BB9	Q96BB9 homo sapien

17	46	51.7	844	12 Q91LD2	Q91LD2 white spot
18	46	51.7	844	12 Q8QZQ0	Q8QZQ0 white spot
19	46	51.7	984	16 Q67280	Q67280 aquilex aeo
20	45	50.6	119	11 Q920E7	Q920E7 mus musculu
21	45	50.6	150	4 Q8NCF2	Q8NCF2 homo sapien
22	45	50.6	169	4 Q8NCF1	Q8NCF1 homo sapien
23	45	50.6	458	4 Q8WVP3	Q8WVP3 homo sapien
24	45	50.6	1852	5 Q812U9	Q812U9 plasmodium
25	45	50.6	4930	16 Q31783	Q31783 bacillus su
26	44	49.4	71	7 Q9GJ67	Q9GJ67 salmo trutt
27	44	49.4	71	7 Q9GJ70	Q9GJ70 salmo trutt
28	44	49.4	71	7 Q9GJ72	Q9GJ72 salmo trutt
29	44	49.4	71	7 Q9GJ69	Q9GJ69 salmo trutt
30	44	49.4	71	7 Q9GJ68	Q9GJ68 salmo trutt
31	44	49.4	71	7 Q9GJ64	Q9GJ64 salmo trutt
32	44	49.4	140	4 Q9UNM8	Q9UNM8 homo sapien
33	44	49.4	472	4 Q8NBZ1	Q8NBZ1 homo sapien
34	44	49.4	472	4 Q91646	Q91646 homo sapien
35	44	49.4	513	10 Q8S658	Q8S658 oryza sativ
36	44	49.4	541	4 Q9Y5X6	Q9Y5X6 homo sapien
37	44	49.4	1104	5 Q812S2	Q812S2 plasmodium
38	44	49.4	1812	5 Q81538	Q81538 plasmodium
39	43	48.3	71	7 Q9GJ66	Q9GJ66 salmo trutt
40	43	48.3	71	7 Q9GJ63	Q9GJ63 salmo trutt
41	43	48.3	71	7 Q9GJ65	Q9GJ65 salmo trutt
42	43	48.3	71	7 Q9GJ62	Q9GJ62 salmo trutt
43	43	48.3	71	7 Q9GJ67	Q9GJ67 brachydanio
44	43	48.3	71	7 Q9XR65	Q9XR65 sigmodon hi
45	43	48.3	134	6 Q9BEB6	Q9BEB6 felis silve

ALIGNMENTS

RESULT 1									
ID	Q9ULB6	PRELIMINARY;	PRT;	95 AA.					
AC	Q9ULB6;								
DT	01-MAY-2000 (TREMBLrel. 13, Created)								
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)								
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)								
DE	Immunoglobulin heavy chain (Fragment).								
GN	VH.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.								
OC	NCBI_TaxID=9606;								
OX	[1]								
RP	SEQUENCE FROM N.A.								
RA	Tange Y., Kayano H.;								
RT	"Human VH gene sequence."								
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AB035268; BAA87067.1; -.								
DR	HSSP; P01772; 2FBA.								
DR	InterPro; IPR007110; IG-like.								
DR	InterPro; IPR003006; IG_MHC.								
DR	InterPro; IPR003596; IG_v.								
DR	Pfam; PF00047; Ig_1.								
DR	SMART; SM00406; Igv_1.								
DR	PROSITE; PS50835; IG_LIKE; 1.								
FT	NON_TER	1	1						
FT	NON_TER	95	95						
SQ	SEQUENCE	95 AA;	10527 MW;	90A8C6D16D22574A	CRC64;				
Query Match									
Best Local Similarity 92.1%; Score 82; DB 4; Length 95;									
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;									
QY	1 NIKQDSEKXYADSVRG 17								
DB	49 NIKQDSEKXYADSVRG 65								

```
RESULT 2
Q9UL90 PRELIMINARY; PRT; 113 AA.
AC Q9UL90;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
  Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
  Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035024; AAD56260.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_1.
DR SMART; SM00406; IGV_1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT TER 113
SQ SEQUENCE 113 AA; 12437 MW; ED57PDD19086D07F CRC64;

Query Match 74.2%; Score 66; DB 4; Length 113;
Best Local Similarity 75.0%; Pred. No. 0.0034;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKDGSSEKYYADSVRG 17
Db 51 IRYDGSNKYYADSVKG 66

RESULT 3
Q9UL93 PRELIMINARY; PRT; 116 AA.
AC Q9UL93;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
  Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
  Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_1.
DR SMART; SM00406; IGV_1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT TER 116
SQ SEQUENCE 116 AA; 12437 MW; ED57PDD19086D07F CRC64;
```

```
SEQ SEQUENCE 116 AA; 12434 MW; 0DA0348154D5C061 CRC64;

Query Match 71.9%; Score 64; DB 4; Length 116;
Best Local Similarity 75.0%; Pred. No. 0.0076;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 IKDGSSEKYYADSVRG 17
Db 50 ISYDGSNKYYADSVKG 65

RESULT 4
Q8WUK1 PRELIMINARY; PRT; 613 AA.
AC Q8WUK1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDA databases.
DR EMBL; BC020240; AAH20240.1; -.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_5.
DR SMART; SM00406; IGV_1.
DR PROSITE; PSS0835; IG_LIKE; 5.
DR PROSITE; PSS00290; IG_MHC; 3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671B315 CRC64;

Query Match 71.9%; Score 64; DB 4; Length 613;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 IKDGSSEKYYADSVRG 17
Db 70 ISYDGSNKYYADSVKG 85

RESULT 5
Q9UL84 PRELIMINARY; PRT; 122 AA.
AC Q9UL84;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
  Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
  Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035030; AAD56266.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003006; IG_MHC.
```


DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 13579 MW; 36054D41366545B8 CRC64;

Query Match
Best Local Similarity 68.5%; Score 61; DB 4; Length 122;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
DB 51 ISNDGSKKYYADSVRG 66

RESULT 6

Q9UL71 PRELIMINARY; PRT; 121 AA.
AC Q9UL71;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9827739; PubMed-9614934;
Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berner S.M., Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
RT Clin. Immunol. Immunopathol. 87:184-192(1998).
RL EMBL; AF035043; AAD56279.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFA5D50736 CRC64;

Query Match
Best Local Similarity 59.6%; Score 53; DB 4; Length 121;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
DB 51 ISDGSSTYYADSVRG 66

RESULT 7

Q9Y509 PRELIMINARY; PRT; 147 AA.
AC Q9Y509;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Vh3 protein (Fragment).
GN Vh3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxId=9606;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE-96071149; PubMed-7475288;
Cao J., Vesic R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C., Lichtenstein A.K., Berenson J.R.;
RT "A CD10-positive subset of malignant cells is identified in multiple myeloma using PCR with patient-specific immunoglobulin gene primers."
Leukemia 9:1948-1953(1995).
RL EMBL; S80860; AAD14339.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1 1
SQ SEQUENCE 147 AA; 15768 MW; 8489FCAA7BC925C CRC64;

Query Match
Best Local Similarity 58.4%; Score 52; DB 4; Length 147;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
DB 51 ISYDSTGYVAGSVRG 66

RESULT 8

Q9GJ71 PRELIMINARY; PRT; 71 AA.
AC Q9GJ71;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN SATR-DAA.
OS Salmo trutta (Brown trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OK NCBI_TaxId=8032;
RN [1]
RP SEQUENCE FROM N.A.
RA Stet R.J.M., Jordan W.C.;
RT "Class II alpha chain in brown trout (Salmo trutta)."
Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ293950; CAC08187.1; -.
DR InterPro; IPR01003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1 1
FT NON_TER 71 71
SQ SEQUENCE 71 AA; 7699 MW; 0E3D9764B397FF72 CRC64;

Query Match
Best Local Similarity 56.2%; Score 50; DB 7; Length 71;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 DGSSEKYYADSVRG 17
DB 14 DGEKYYADSVRG 26

RESULT 9

Q8NC16 PRELIMINARY; PRT; 493 AA.
AC Q8NC16;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein FLN90170.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Isegai T., Ota T., Nishikawa T., Hayashi K., Otauki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "MEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074651; BAC11114.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG cl.
DR InterPro; IPR003066; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; IG; 4.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGC1; 2.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
DR Hypothetical protein.
SQ SEQUENCE 493 AA; 53224 MW; 12ECD7E094777101 CRC64;

Query Match 56.2%; Score 50; DB 4; Length 493;
Best Local Similarity 61.5%; Pred. No. 8.2;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 DGESEKYADSVRG 17
Db 73 DGTXYSDSVKG 85
||:|||||:|
|:|:|:|:|:|

RESULT 10
Q9UL72 PRELIMINARY; PRT; 118 AA.
AC Q9UL72;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalle N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035042; AAD56278.1; -.
DR HSP; P01772; 2F84.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003066; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

Query Match 53.9%; Score 48; DB 4; Length 118;
Best Local Similarity 47.1%; Pred. No. 3.5;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NIKDGESEKYADSVRG 17

```

```

Db 49 SVTYSGSSYADSVRG 65
||:|||||:|
|:|:|:|:|:|

RESULT 11
Q92HB9 PRELIMINARY; PRT; 593 AA.
AC Q92HB9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Penicillin-binding protein.
GN PPP1 OR RC0852.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Oyata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cosart P., Weissenbach J., Claverie J.-M.,
RA Rault D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL; AB008640; AAL03390.1; -.
DR InterPro; IPR005311; PBP_dimer.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF03717; PBP_dimer; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR Complete proteome.
SQ SEQUENCE 593 AA; 66518 MW; F1CA3B85086C66F1 CRC64;

Query Match 53.9%; Score 48; DB 16; Length 593;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NIKDGESEKYADSVRG 17
Db 194 NIKSGIEKYDKLKG 210
||:|||||:|
|:|:|:|:|:|

RESULT 12
Q9ZCY6 PRELIMINARY; PRT; 594 AA.
AC Q9ZCY6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Penicillin-binding protein (PBP1).
GN PP565.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid B;
RX MEDLINE=99039499; PubMed=9823893;
RA Anderson S.G.E., Zomrodipour A., Anderson J.O.,
RA Scharfetz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kuriland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 386:133-140(1998).
DR EMBL; AJ235272; CAAL5013.1; -.
DR InterPro; IPR005311; PBP_dimer.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF03717; PBP_dimer; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR Complete proteome.
SQ SEQUENCE 594 AA; 67195 MW; 669AB06B4A2BAC CRC64;

```

	Query Match	53.9%	Score 48;	DB 16;	Length 594;	
	Best Local Similarity	52.9%	Pred. No. 22;			
	Matches	9;	Conservative	3;	Mismatches	5; Indels
						Gaps
Dd						
	1 NIKODGSEKRYADSVRG 17					
	:::: :: :: :					
	194 NIGKSIEKYDNKLKG 210					
RESULT 13						
Q9VZF9	PRELIMINARY;	PRT;	247 AA.			
AC Q9VZF9	OBSZQ3;					
DT 01-MAY-2000	(TREMBLrel. 13. Created)					
DT 01-OCT-2002	(TREMBLrel. 22. Last sequence update)					
DT 01-OCT-2002	(TREMBlrel. 22. Last annotation update)					
DE CG1259	protein (KE69226p).					
GN CG1259.						
OS Drosophila melanogaster (Fruit fly).						
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;						
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;						
OC Epiphydoidea; Drosophilidae; Drosophila.						
OX NCBI_taxonomy:7227;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC STRAIN=Berkelley;						
RX MEDLINE=20196006; PubMed=10731132.						
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,						
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,						
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,						
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,						
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,						
RA Wan K.H., Doyle C., Baxter E.C., Helt G., Nelson C.R., Miklos G.L.G.,						
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,						
RA Ballen R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,						
RA Beeson K.Y., Benos P.V., Bertan B.P., Bhattacharya D., Bolshakov S.,						
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brodtier P.,						
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,						
RA Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,						
RA de Pablos B., Delcher A., Deng Z., Meyers A.D., Dew I., Dietz S.M.,						
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,						
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,						
RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glaesner K.,						
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,						
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,						
RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibeswari C.,						
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Keithum K.A.,						
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,						
RA Laake P., Lei Y., Levitsky A.C., Li J.Z., Liang Y., Lin X.,						
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,						
RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Mosheret A.,						
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.T.,						
RA Neilson S.R., Nelson K.A., Nixon K., Nuskern D.R., Paclel J.M.,						
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,						
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,						
RA Shue B.C., Sliden-Klimos I., Simpson M., Skupski M.P., Smith T.,						
RA Spier E.B., Spradling A.C., Stapleton M., Strong R., Sun E.,						
RA Svirskaas R., Tector C., Turner R., Ventzer E., Wang A.H., Wang X.,						
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,						
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,						
RA Ye J., Yeh R.-F., Zavertil J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,						
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,						
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;						
"The genome sequence of Drosophila melanogaster.";						
Science 287:2185-2195(2000).						
[2]						
SEQUENCE FROM N.A.						
Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,						
Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,						
Batzon J., An H., Baldwin D., Batzon J., Beeson K.Y., Busam D.A.,						
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,						
Dodon K., Dorsett V., Doup L.E., Doyle C., Drenek D., Farfan D.,						

```

RA Ferriser S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoakins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Meshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacel J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pitman G.S., Puri V., Richards S., Scheeler F.,
RA Stepleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Miara S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hirdecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tuzy J.L., Bergman C., Bertram B., Carlson J.W., Celinker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnak D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarini H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungal C.J., Nunoo J., Pacel J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003480; AAP47864.2; -
DR EMBL; AY070590; AAL48061.1; -
DR FlyBase; FBGN0035513; CG1259.
DR InterPro; IPR000618; Insect_cuticle.
DR Pfam; PF00379; Chitin_bind_4; 1.
DR PRINTS; PR00947; CUTICLE.
DR PROSITE; PS00233; CUTICLE; 1.
SQ SEQUENCE 247 AA; 24279 MW; 3C46B2995A7AAD1 CRC64;

Query Match 53.4%; Score 47.5; DB 5; Length 247;
Best Local Similarity 52.6%; Pred. No. 9.7;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1.

CY 2 IKODGSEK--YYADSVRG 17
Db |::|::|::|::|::|
179 IEPPGSRIVSYADSLNG 197

RESULT 14
O9F5K2 PRELIMINARY; PRT; 342 AA.
AC O9F5K2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Probable ABC-transporter binding protein.
GN ORF342.
OS Rhizobium tropici.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
CX NCBI_TaxId=398;
XN [1]
XP SEQUENCE FROM N.A.
RC STRAIN=CIA7 899;
```

RA Becker B.U., Neumann-Silkow F., Mueller P.;
 RT "ORF 342, a potential binding protein of an ABC transporter in
 Rhizobium tropici CIAT 899."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517883; AAC31660.1; -.
 DR InterPro; IPR006059; SBP_bac_1.
 DR InterPro; IPR001188; Sperm/putr-bndg.
 DR Pfam; PF01547; SBP_bac_1; 1.
 DR PRINTS; PR00909; SPERMNDNDG.
 SQ SEQUENCE 342 AA; 3797 MW; CCBBD9D01767628 CRC64;

Query Match 52.8%; Score 47; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 17;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 GSEKTYADSVRG 17
 |:|||||:|:|
 Db 294 GAERYSDDMKG 305

RESULT 15

ID Q987Q4 PRELIMINARY; PRT; 360 AA.
 AC Q987Q4;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE ABC transporter, binding protein.
 GN MNR6964.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099; PubMed=11214968;
 RX MEDLINE=21082930;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M.,
 RA Wochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti."
 RT DNA Res. 7:331-338(2000).
 RL EMBL; AF003010; BAB53149.1; -.
 DR InterPro; IPR006059; SBP_bac_1.
 DR InterPro; IPR001188; Sperm/putr-bndg.
 DR Pfam; PF01547; SBP_bac_1; 1.
 DR PRINTS; PR00909; SPERMNDNDG.
 KW Complete proteome.
 SQ SEQUENCE 360 AA; 39917 MW; 570A061F804F6EDB CRC64;

Query Match 52.8%; Score 47; DB 16; Length 360;
 Best Local Similarity 58.3%; Pred. No. 18;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 GSEKTYADSVRG 17
 |:|||||:|:|
 Db 312 GAERYSDDMKG 323

Search completed: January 12, 2004, 07:01:58
 Job time : 38.5312 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:09 ; Search time 41.1719 Seconds
(without alignments)
65.539 Million cell updates/sec

Title: US-09-829-495-62

Sequence: 1 NIKQDGSEKYYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

1: A_GeneSeq_19Jum03:*

2: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1980.DAT:*

3: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1981.DAT:*

4: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1982.DAT:*

5: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1983.DAT:*

6: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1984.DAT:*

7: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1986.DAT:*

8: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1987.DAT:*

9: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1988.DAT:*

10: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1990.DAT:*

11: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1990.DAT:*

12: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1991.DAT:*

13: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1992.DAT:*

14: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1993.DAT:*

15: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1994.DAT:*

16: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1995.DAT:*

17: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1996.DAT:*

18: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1997.DAT:*

19: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1998.DAT:*

20: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA1999.DAT:*

21: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA2000.DAT:*

22: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA2001.DAT:*

23: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA2002.DAT:*

24: /SIDS1/gcgcatc/geneeq/genseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	89	100.0	17	22	AA861294	Anti-TANGO 268 scf
2	89	100.0	17	24	ABU11260	Human TANGO 268 v1
3	82	92.1	17	23	AAO17611	Human FapAlp1pha sp.
4	82	92.1	18	15	AA852066	Heavy Chain variata
5	82	92.1	98	16	AA72074	DP54 VH region. 1
6	82	92.1	98	21	AA40138	Anti-H1L2 antibody
7	82	92.1	98	21	AA840140	Anti-H1L2 antibody
8	82	92.1	98	23	ABG91901	Human antibody fr
9	82	92.1	98	23	ABG78210	Human fv molecule

10	82	92.1	117	16	AA86301	Human immunoglobul
11	82	92.1	125	22	AA862776	Human HIV-1 monocl
12	82	92.1	134	23	AA014069	Human anti-FAPalph
13	82	92.1	241	23	ABP45878	Human Blys binding
14	82	92.1	247	23	ABP45459	Human Blys binding
15	82	92.1	251	23	ABP45583	Human Blys binding
16	82	92.1	253	23	ABP44992	Human Blys binding
17	82	92.1	253	23	ABP44996	Human Blys binding
18	82	92.1	262	23	AA014058	Human Blys binding
19	82	92.1	271	20	AA890180	Human clone A6 fus
20	82	92.1	312	20	AA883333	Single chain Apo-2
21	82	92.1	312	23	AB809604	Amino acid sequenc
22	82	92.1	312	24	ABG74385	Single chain antib
23	78	87.6	17	23	AAU75740	Anti-(MCP)-1 antib
24	78	87.6	122	23	AAU75737	AAV93 anti-(MCP)-
25	76	85.4	264	23	AAU75152	Protein for affini
26	76	85.4	265	23	AAU75149	Amino acid sequenc
27	76	85.4	265	23	AAU75153	Protein for clone
28	76	85.4	265	23	AAU75154	Protein for clone
29	75	84.3	17	20	AAV05049	Tumour antigen ant
30	75	84.3	250	23	ABP45308	Human Blys binding
31	75	84.3	250	23	ABP45475	Human Blys binding
32	74	83.1	122	23	AAU75745	Human Blys binding
33	74	83.1	254	23	ABP45500	Human Blys binding
34	73	82.0	17	20	AAV05055	Tumour antigen ant
35	73	82.0	254	23	ABP45339	Human Blys binding
36	72	80.9	17	21	AA839884	Anti-htl2 antibod
37	71	79.8	17	18	AAW44547	CDR #2 of r101-2 h
38	71	79.8	130	18	AAW44537	Immunoglobulin r10
39	69	77.5	17	21	AA839883	Anti-htl2 antibod
40	69	77.5	116	16	AA866340	Human immunoglobul
41	67	75.3	17	21	AA839882	Anti-htl2 antibod
42	67	75.3	117	22	AAG33629	Human anti-Rh(D) a
43	67	75.3	123	22	AA862778	Human HIV-1 monocl
44	67	75.3	249	23	ABP45806	Human Blys binding
45	66	74.2	17	21	AA839467	Anti-htl-12 antibo

ALIGNMENTS

RESULT 1	
AAB61294	
ID	AAB61294 standard; Peptide; 17 AA

AC	AAB61294;
XX	
DT	04-APR-2001 (first entry)

DE Anti-TANGO 268 scFv CDR, SEQ ID NO: 62.

KM Haman; tttbody; scfv; cdr; complementarity determining region;
KM TAMGO; 266; cartian; cerebroprotective; cytosolic; anticoagulant;
KM thrombolytic; antiarteriosclerosis; haemostatic; glycoprotein VI; GPVI;
KM platelet membrane glycoprotein receptor; bleeding disorder;
KM blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;
KM leukaemia; cardiovascular disease; immunological disease; liver disorder
KM cancer.

OS	Homo sapiens.
XX	
PN	W0200100810-A1.
XX	
PD	04-JAN-2001.
XX	
PF	30-JUN-2000; 2000WO-US18152.
XX	
PR	30-JUN-1999; 99US-0345468.
PR	06-DEC-1999; 99US-0454824.
PR	14-FEB-2000; 2000US-0503387.
XX	
PA	
XX	(MILL-) MILLENNIUM PHARM INC

PI Busfield SJ, Villelail J, Jandrot-Perrus M, Vainchenker W, Gili DS;
PI Qian DM, Kingsbury G;
XX
XX WPI; 2001-080877/09.
XX
XX New genes encoding human platelet-expressed collagen receptor,
PT glycoprotein VI, and its modulators, useful for preventing, treating
PT and diagnosing hemorrhagic disorders, thrombotic diseases and
PT immunological disorders -
XX
XX Claim 31; Page 102; 227pp; English.
XX
XX The present sequence is given in a specification relating to an isolated
CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
CC and polypeptides and their modulators, e.g. antisense nucleic acids,
CC ribozymes and antibodies, are useful for preventing, treating and
CC diagnosing disorders associated with aberrant expression or activity of
CC GPVI. These disorders include bleeding disorders
CC (e.g. thrombocytopenia), blood vessel injury, thrombotic disorders
CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC infarction), immunological diseases (e.g. platelet disorder) and
CC embryonic liver disorders. Preferably they are used to prevent acture
CC cardiac ischaemia following angioplasty and metastatic cancers,
CC especially of the colon and liver.
XX
XX Sequence 17 AA;
SQ
Query Match 100.0%; Score 89; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NIKQDSEKXYADSVRG 17
Db 1 NIKQDSEKXYADSVRG 17
RESULT 2
ABU11260 standard; Peptide; 17 AA.
XX
XX ABU11260;
AC
DT 06-FEB-2003 (first entry)
XX
XX Human TANGO 268 VHCDR2 Peptide #3.
DE
XX
XX Human; mouse; variable heavy; VH; antigen; cancer;
KM complementarity determining region; TANGO 268; glycoprotein VI; GPVI;
KM TANGO 268; extracellular matrix; collagen; platelet release;
KM proliferation; migration; embryogenesis; inflammation; thrombosis;
KM degradation; thrombocytopenia; antibody; thrombotic disorder;
KM cerebral vascular disease; stroke; ischaemia; venous thromboembolism;
KM leg swelling; pain; ulceration; pulmonary embolism; coronary disease;
KM cardiovascular disease; angina pectoris; myocardial infarction;
KM coronary restenosis; atherosclerosis; immunological disorder;
KM developmental disorder; embryonic disorder; liver disorder;
KM cerebral vascular disease; venous thromboembolism disease.
XX
XX Homo sapiens.
OS
XX
XX WO200280968-A1.
PN
XX 17-OCT-2002.
PD
XX 09-APR-2002; 2002WO-US11122.
PF
XX 09-APR-2001; 2001US-0829495.
PR
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX

PI Busfield SJ, Villelail J, Jandrot-Perrus M, Vainchenker W, Gili DS;
PI Qian DM, Kingsbury G;
XX
XX WPI; 2003-058477/05.
XX
XX Novel substantially purified antibody immunospecifically binding to
PT TANGO 268 antigen, useful for treating bleeding disorders such as
PT thrombocytopenia, stroke, ischaemia, pulmonary embolism, atherosclerosis
PT
XX
XX Claim 7; Page 111; 236pp; English.
PS
XX
XX This invention relates to a novel purified antibody comprising a
CC variable heavy (VH) complementarity determining region (CDR)1, VH CDR2
CC or VH CDR3; or variable light (VL) CDR1, VL CDR2 or VL CDR3, and
CC immunospecifically binding to a TANGO 268 (also referred as glycoprotein
CC VI (GPVI)) antigen. The antibodies of the invention act to decrease or
CC block TANGO 268 binding to extracellular matrix components, or as a
CC collagen or platelet release and aggregation blocker. The antibodies of
CC the invention are useful for modulating proliferation, migration,
CC morphology, differentiation and/or function of megakaryocytes and
CC platelets, including during development e.g. embryogenesis, modulating
CC leukocyte-platelet and platelet-endothelium interactions in
CC inflammation and/or thrombosis, and modulating platelet aggregation and
CC degradation. They are also useful for modulating disorders associated
CC with abnormal or aberrant megakaryocyte and/or platelet proliferation,
CC migration, morphology, differentiation and/or function, e.g. bleeding
CC disorders such as thrombocytopenia. Other diseases which may be
CC modulated by these antibodies are thrombotic disorders, cerebral
CC vascular diseases (e.g. stroke and ischaemia) venous thromboembolism
CC diseases (e.g. diseases involving leg swelling, pain and ulceration,
CC pulmonary embolism, etc); coronary diseases (e.g. cardiovascular
CC diseases including angina pectoris, myocardial infarction, coronary
CC restenosis, atherosclerosis, etc); immunological disorders,
CC developmental disorders, embryonic disorders, liver disorders, cerebral
CC vascular diseases, venous thromboembolism disease, coronary diseases,
CC and metastatic cancers. The antibodies of the invention only causes a
CC transient decrease in platelet counts, platelet aggregation, and/or
CC platelet activation and so have some advantages over prior art
CC methods. The present sequence represents a peptide sequence used to
CC generate the antibodies of the invention.
XX
XX Sequence 17 AA;
SQ
Query Match 100.0%; Score 89; DB 24; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NIKQDSEKXYADSVRG 17
Db 1 NIKQDSEKXYADSVRG 17
RESULT 3
AAO17611 standard; Peptide; 17 AA.
ID AAO17611
XX
XX AAO17611;
AC
DT 08-AUG-2002 (first entry)
XX
XX Human PAPAlpha specific VL region from VH50 CDR2 peptide.
DE
XX
XX Human; PAPAlpha; fibroblast activating protein alpha; antibody; Ab;
KM gene therapy; cancer; wound healing; inflammation; cytostatic.
XX
XX Homo sapiens.
OS
XX
XX WO200168708-A2.
PN
XX 20-SEP-2001.
PD
XX 16-MAR-2001; 2001WO-EP04716.
PF

XX (NICH-) NICHOLS INST DIAGNOSTICS.
 PA McIlachlan SM, Rapoport B;
 PI WPI; 1995-139383/18.
 DR N-PSDB; AAQ89332.
 XX
 PT Graves' ophthalmopathy-associated monoclonal antibody - produced
 by molecular cloning of immunoglobulin genes by PCR
 XX
 PS Disclosure, Page 72; 94pp; English.
 CC L- and H-chain DNA was amplified by PCR from Graves' orbital
 CC tissue and clones encoding autoimmunity-associated immunoglobulin
 CC fragments were obtained. 2/15 clones of H chain (IgG1) genes
 CC showed homology to the germ-line gene D554 (AA089332), which encodes
 CC the protein given in AAR72074. The DNA (AA089333) and corresp. amino
 CC acid (AAR72075) sequences of the VH region of a representative clone,
 CC OF7H1.19, are provided.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 SQ Sequence 98 AA;
 Query Match 92.1%; Score 82; DB 16; Length 98;
 Best Local Similarity 88.2%; Pred. No. 3.4e-06;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

PS Claim 75; Page 122; 377pp; English.
 CC This invention relates to a new human antibody specific for human
 CC interleukin-12 (IL-12). The invention also includes antigen binding
 CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
 CC anti-IL-12 antibody heavy and light chain complementarity determining
 CC region (CDR) amino acid sequences, and also includes variable region
 CC amino acid sequences. Other variable region amino acid sequences are
 CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
 CC represent anti-IL-12 CDR3 related amino acid sequences. AAB39772-B40063
 CC represent other CDR sequences. Light chain CDR3 consensus sequences are
 CC given in AAB40064-B40067. Primers used in the identification and
 CC construction of the antibodies of the invention are given in
 CC AAC61062-C61071. The antibody of the invention is a neutralizing
 CC antibody and has antirheumatic; antiarthritic; anti-sclerotic;
 CC antiinflammatory; neuroprotective; antipsoriatic; antiasthmatic;
 CC cardiant; antiparasitic; antibacterial and immunosuppressive activity.
 CC The antibodies or antigen-binding fragments are useful in the treatment
 CC of disorders associated with detrimental release of human IL-12,
 CC especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
 CC They can also be used in the manufacture of a pharmaceutical composition
 CC to treat human IL-12 disorders.
 CC
 SQ Sequence 98 AA;
 Query Match 92.1%; Score 82; DB 21; Length 98;
 Best Local Similarity 88.2%; Pred. No. 3.4e-06;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODGSEKYYADSVRG 17
 DB 50 NIKODGSEKYYVDSVKG 66

OY 1 NIKODGSEKYYADSVRG 17
 DB 50 NIKODGSEKYYVDSVKG 66

RESULT 6
 AAB40138
 ID AAB40138 standard; Protein; 98 AA.
 XX
 AC AAB40138;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Anti-HIL12 antibody H chain V region amino acid sequence SEQ ID 664.
 XX
 KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
 KW complementarity determining region; CDR; antirheumatic; antiarthritic;
 KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
 KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 PN WO200056772-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 24-MAR-2000; 2000WO-US07946.
 XX
 PR 25-MAR-1999; 99US-0126603.
 XX
 PA (BADI) BASF AG.
 PA (GEMY) GENETICS INST INC.
 XX
 PI Safeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M,
 PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A,
 PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR,
 PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
 XX
 DR WPI; 2000-638250/61.
 XX
 PT New human antibody specific for human interleukin-12 (IL-12) used to
 PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
 PT disease and multiple sclerosis -

RESULT 7
 AAB40140
 ID AAB40140 standard; Protein; 98 AA.
 XX
 AC AAB40140;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Anti-HIL12 antibody H chain V region amino acid sequence SEQ ID 666.
 XX
 KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
 KW complementarity determining region; CDR; antirheumatic; antiarthritic;
 KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
 KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 PN WO200056772-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 24-MAR-2000; 2000WO-US07946.
 XX
 PR 25-MAR-1999; 99US-0126603.
 XX
 PA (BADI) BASF AG.
 PA (GEMY) GENETICS INST INC.
 XX
 PI Safeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M,
 PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A,
 PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR,
 PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
 XX
 DR WPI; 2000-638250/61.
 XX
 PT New human antibody specific for human interleukin-12 (IL-12) used to
 PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
 PT disease and multiple sclerosis -

XX Claim 75, Page 122, 377pp; English.

PS

XX This invention relates to a new human antibody specific for human

CC interleukin-12 (IL-12). The invention also includes antigen binding

CC portions that bind to IL-12. Sequences ABB39485-B39516 represent human

CC anti-IL-12 antibody heavy and light chain complementarity determining

CC region (CDR) amino acid sequences, and also includes variable region

CC amino acid sequences. Other variable region amino acid sequences are

CC given in ABB39517-B39560 and ABB40068-B40149. Sequences ABB39561-B39771

CC represent anti-IL-12 CDR3 related amino acid sequences. ABB39772-B40063

CC represent other CDR sequences. Light chain CDR3 consensus sequences are

CC given in ABB40064-B40067. Primers used in the identification and

CC construction of the antibodies of the invention are given in

CC AAC61062-C61071. The antibody of the invention is a neutralising

CC antibody and has antitumour, antirheumatic, antileukemic,

CC antiinflammatory, neuroprotective, antipsoriatic, antistatic,

CC cardiant, antiparasitic, antibacterial and immunosuppressive activity.

CC The antibodies or antigen-binding fragments are useful in the treatment

CC of disorders associated with detrimental release of human IL-12,

CC especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.

CC They can also be used in the manufacture of a pharmaceutical composition

CC to treat human IL-12 disorders.

CC

SQ Sequence 98 AA;

Query Match 92.1%; Score 82; DB 21; Length 98;

Best Local Similarity 88.2%; Pred. No. 3.4e-06;

Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKRYADSVRG 17

|||||||

Db 50 NIKODSEKRYADSVRG 66

RESULT 8

ABG91901

ID ABG91901 standard; Protein; 98 AA.

XX

AC ABG91901;

XX

DT 04-DEC-2002 (first entry)

XX

DE Human antibody fragment #85.

XX

KW Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;

XX metastasis; hypervariable region; autoimmune disease; thrombosis;

XX restenosis; leukaemia; inflammatory disease; cardiovascular disease;

XX myocardial infarction; retinopathic disease; abnormal platelet function;

XX sulphated tyrosine-dependent protein-protein interaction.

XX

OS Homo sapiens.

XX

PN WO200253700-A2.

XX

PD 11-JUL-2002.

XX

PF 31-DEC-2001; 2001WO-US49442.

XX

PR 29-DEC-2000; 2000US-25848P.

XX

PR 29-DEC-2000; 2000US-0751181.

XX

PA (BIOT-) BIO-TECHNOLOGY GEN CORP.

XX

PI Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H,

XX Szanton E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;

XX WPI; 2002-674776/72.

XX

PT Novel isolated epitope present on cancer cells and important in

XX physiological phenomena such as cell rolling, metastasis and

XX inflammation, for treating autoimmune, inflammatory or cardiovascular

XX diseases, and cancer

PT

XX Disclosure; Page 268; 310pp; English.

PS

XX

CC The invention relates to an isolated epitope present on cancer cells and

CC important in physiological phenomena such as cell rolling, metastasis and

CC inflammation, where the epitope is capable of being bound by an antibody,

CC its antigen-binding fragment or its complex comprising at least one

CC antibody or its binding fragment having a first hypervariable region. The

CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune

CC disease, thrombosis, restenosis, metastasis, growth and/or replication of

CC tumour or leukaemia cells, increase in number of tumour or leukaemia

CC cells in a patient, cell-cell, cell-matrix, platelet-matrix,

CC platelet-platelet and/or cell-platelet adhesion or aggregation, for

CC increasing mortality of tumour or leukaemia cells, for increasing the

CC susceptibility of diseased cells to damage by anti-disease, anti-cancer

CC or anti-leukaemia agents, or for decreasing the number of tumour or

CC leukaemia cells in a patient, or in the manufacture of a medicament for

CC the above mentioned purposes. The epitopes are useful for diagnosing and

CC treating diseases such as cancer, leukaemia, autoimmune diseases,

CC inflammatory diseases, cardiovascular diseases such as myocardial

CC infarction, retinopathic diseases and other diseases mediated by abnormal

CC platelet function and diseases caused by sulphated tyrosine-dependent

CC protein-protein interactions. This sequence represents a human antibody

CC fragment of the invention.

CC

SQ Sequence 98 AA;

Query Match 92.1%; Score 82; DB 23; Length 98;

Best Local Similarity 88.2%; Pred. No. 3.4e-06;

Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKRYADSVRG 17

|||||||

Db 50 NIKODSEKRYADSVRG 66

RESULT 9

ABG78210

ID ABG78210 standard; Protein; 98 AA.

XX

AC ABG78210;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human Fv molecule hypervariable region related peptide #85.

XX

KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;

XX disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;

XX lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.

XX

OS Homo sapiens.

XX

PN WO200259264-A2.

XX

PD 01-AUG-2002.

XX

PF 31-DEC-2001; 2001WO-US49440.

XX

PR 29-DEC-2000; 2000US-0751181.

XX

PA (BIOT-) BIO-TECHNOLOGY GEN CORP.

XX

PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;

XX Plaksin D, Peretz T;

XX WPI; 2002-619166/66.

XX

PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct

XX or fragment, or construct of fragment with enhanced binding

XX characteristics so as to selectively bind target cell in favour of other

XX cells

PT

XX Claim 13; Page 191-192; 232pp; English.

XX CC The invention relates to a peptide or polypeptide comprising an Fv
CC molecule, a construct or fragments or a construct of a fragment with
CC enhanced binding characteristics which selectively and/or specifically
CC binds to a target cell in favour of other cells, where binding is
CC primarily determined by a first hypervariable region and Fv is a single
CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC association with or attached, coupled, combined, linked or fused to a
CC pharmaceutical agent, is useful in the manufacture of a medicament, where
CC the medicament has activity against a diseased cell, preferably a cancer
CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC composition for use in inhibiting the growth of a diseased or cancer
CC cell. This sequence represents a human Fv molecule hypervariable region
CC related peptide of the invention.
XX
SQ Sequence 98 AA;
Query Match 92.1%; Score 82; DB 23; Length 98;
Best Local Similarity 88.2%; Pred. No. 3.4e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKDGSSEKYYADSVRG 17
DB 50 NIKDGSSEKYYADSVRG 66
RESULT 10
AAR6301
ID AAR6301 standard; Protein; 117 AA.
XX
AC AAR6301;
XX
DT 25-MAR-2003 (updated)
DT 02-AUG-1995 (first entry)
XX
DE Human immunoglobulin variable heavy chain #7.
XX
KW Primer; PCR; amplify; human; immunoglobulin; variable; heavy chain;
KW cosmid; placenta; vector; pJb81; E.coli; mammalian.
XX
OS Homo sapiens.
XX
PN WO9426895-A1.
XX
PD 24-NOV-1994.
XX
PF 10-MAY-1993; 93WO-JP00603.
XX
PR 10-MAY-1993; 93WO-JP00603.
XX
PA (NIBS) JAPAN TOBACCO INC.
XX
PI Honjo T, Matsuda F;
XX
DR WPI: 1995-006791/01.
DR N-PSDB; AAO78946.
XX
PT DNA fragment comprising human immunoglobulin Vh genes - for the
PT production of human immunoglobulin in mammalian hosts
XX
PS Claim 17, Page 40-41; 130pp; Japanese.
XX
CC Protein sequences (AAR6295-51) are novel human immunoglobulin heavy
CC chain sequences encoded by novel isolated genes. The genes
CC (AAO78939-79002) were isolated and cloned from a series of cosmid
CC constructs: Y202; Y103; Y21; Y6; Y24; 3-31; M84; M18 and M31, by PCR
CC amplification using primers AAO78917-38. The genes are subdivided into 5
CC families of Vh genes. The fragments cover a region of 800 kb. The DNA
CC fragments were isolated from high molecular weight DNA from human
CC placenta. The DNA was partially digested with TaqI restriction enzyme.
CC The fragments were separated by gel electrophoresis and 35-45 kb fractions

CC were collected. The fragments were ligated with ClaI-digested cosmid
CC vector pJb81. The ligation products were in vitro packed and infected
CC into E.coli 490A. The fragments were then subcloned by colony
CC hybridisation. The Vh genes and the DNA fragments encoding them are
CC useful in producing human immunoglobulin in mammalian hosts.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 117 AA;
Query Match 92.1%; Score 82; DB 16; Length 117;
Best Local Similarity 88.2%; Pred. No. 4.2e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKDGSSEKYYADSVRG 17
DB 69 NIKDGSSEKYYADSVRG 85
RESULT 11
AAB62776
ID AAB62776 standard; Protein; 125 AA.
XX
AC AAB62776;
XX
DT 03-APR-2001 (first entry)
DT
XX
DE Human HIV-1 monoclonal antibody SEQ ID NO: 75.
XX
KW Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;
KW envelope glycoprotein; gp120; diagnosis.
XX
OS Homo sapiens.
XX
PN WO200100678-A1.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-US17327.
XX
PR 30-JUN-1999; 99US-0141701.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Watkins BA, Reitz MS;
XX
PN WPI: 2001-112438/12.
XX
PD N-PSDB; AAF29077.
XX
PF Novel human monoclonal antibody immunoreactive with human
PT immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
PT in biological sample and providing passive immunotherapy to HIV-1
PT infected mammal -
XX
PS Claim 1, Page 69-70; 81pp; English.
XX
CC The present invention provides the protein and coding sequences for the
CC variable regions of human monoclonal antibodies which are immunoreactive
CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
CC These can be used in diagnosis and therapy of HIV-1 infection.
XX
SQ Sequence 125 AA;
Query Match 92.1%; Score 82; DB 22; Length 125;
Best Local Similarity 88.2%; Pred. No. 4.5e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKDGSSEKYYADSVRG 17
DB 51 NIKDGSSEKYYADSVRG 67
RESULT 12
AAO14049

ID		AA014049 standard; Protein; 134 AA.
AC		AA014049;
XX		
DT	08-AUG-2002	(first entry)
DE		Human anti-FAPalpha antibody fragment VH50.
KM		Human; FAPalpha; fibroblast activating protein alpha; antibody; Ab; gene therapy; cancer; wound healing; inflammation; cytostatic.
OS		Homo sapiens.
PN		WO200168708-A2.
PD	20-SEP-2001.	
PF	16-MAR-2001; 2001WO-EP04716.	
PR	17-MAR-2000; 2000DE-1013286.	
RR	11-SEP-2000; 2000GB-0022216.	
PA	(BOEH) BOEHRINGER INGELHEIM PHARMA KG.	
PI	Park J, Garin-Chesa P, Pfizenmaier K, Moosmayer D, Weismann M, Schmidt A;	
DR	WP1; 2002-041180/05.	
DR	N-PSDB; AAK98398.	
PT	New human humanised antibody that specifically binds to fibroblasts activating protein alpha, useful for treating cancer or tumor, and for imaging tumors associated with activated stromal fibroblasts, e.g. lung or breast cancer -	
PS	Claim 16; Page 43; 109pp; English.	
CC	The present invention relates to a human or humanised antibody (Ab) which specifically binds to fibroblast activating protein alpha (FAPalpha). The antibodies are useful for preparing a composition for the treatment of cancer, and for imaging tumours associated with activated stromal fibroblasts, such as colorectal cancer, non-small-cell lung cancer, breast cancer, head and neck cancer, ovarian cancer, lung cancer, bladder cancer, pancreatic cancer and metastatic brain cancer, and diseases associated with the same, such as inflammation and wound healing. The present sequence is a polypeptide described in the exemplification of the invention.	
SQ	Sequence	134 AA;
Query Match	92.1%; Score 82; DB 23; Length 134;	
Best Local Similarity	88.2%; Pred. No. 4.9e-06;	
Matches	15; Conservative	1; Mismatches 1; Indels 0; Gaps 0
OY	1 NIKDGSSEKYVDSVRG 17	
Dd	50 NIKDGSSEKYYVDYVG 66	
RESULT 13		
ID	ABP45878 standard; Protein; 241 AA.	
AC	ABP45878;	
DT	19-AUG-2002 (first entry)	
DE	Human Blys binding scFv SEQ ID 1889.	
KM	Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic; tumor necrosis factor; B cell proliferation; B cell differentiation; immunosuppressive; immunostimulant; immunomodulatory; antirheumatic; antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;	

KW		systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KV		common variable immunodeficiency; acquired immunodeficiency syndrome.
XX	Homo sapiens.	
OS	WO200202641-A1.	
PX	10-JAN-2002.	
XX	15-JUN-2001, 2001WO-US19110.	
PF	16-JUN-2000; 2000US-212210P.	
XX	PR 17-OCT-2000; 2000US-240816P.	
PR	16-MAR-2001; 2001US-276248P.	
FR	21-MAR-2001; 2001US-277379P.	
PR	25-MAY-2001; 2001US-293499P.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	(CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.	
XX	Ruben SW, Barash SC, Choi GH, Vaughan T, Hilbert D,	
EI	WPI; 2002-114799/15.	
DR		
PT	Antibodies against B lymphocyte stimulating polypeptides, useful for the diagnosis and treatment of cancers and immune disorders - Claim 1; Page 2655-2656; 3148pp; English.	
PS	This invention describes novel antibodies that immuno-specifically bind to B lymphocyte stimulator (BLyS) polypeptides. BLyS is a member of the tumour necrosis factor (TNF) super family and induces B cell proliferation and differentiation. The antibodies of the invention have cytostatic, immunosuppressive, immunostimulant, immunomodulatory, antirheumatic and antiAIDS activity and can be used in vaccines to inhibit the expression and activity of BLyS. The antibodies bind to BLyS and so may be used to detect and quantitate the presence of BLyS in biological samples and may be used in this way to diagnose disease associated with aberrant expression of BLyS. They may also be administered to treat diseases associated with aberrant BLyS expression and activity such as cancer, immune, and autoimmune disorders and diseases, e.g., systemic lupus erythematosus, rheumatoid arthritis, immunodeficiency (e.g., common variable immunodeficiency (CVID)) and acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP47228 represent the antibodies and fragments of the antibodies described in the method of the invention.	
CC		
CC		
CC		
SC	Sequence 241 AA;	
BQ		
Query Match	92.1%; Score 82; DB 23; Length 241; Best Local Similarity 88.2%; Pred. No. 9.6e-06;	
Matches	15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
DY	1 NIKODGSEKRYADSVRG 17 :	
DB	50 NIKODGSEKYYVDSDVGK 66	
ID	ABP45459 standard; Protein, 247 AA.	
AC	ABP45459;	
DT	19-AUG-2002 (first entry)	
DE	Human BLyS binding scFv SEQ ID 1470.	
XX		
XX	Blvs; B lymphocyte stimulator; TNF superfamily; human; cytotoxic; tumour necrosis factor; B cell proliferation; B cell differentiation; immunosuppressive; immunostimulant; immunomodulatory; antirheumatic; antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency; systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDs;	

KM common variable immunodeficiency; acquired immunodeficiency syndrome.
XX Homo sapiens.
OS
XX MO200202641-A1.
PN
XX 10-JAN-2002.
PD
XX 15-JUN-2001; 2001WO-US19110.
PF
XX 16-JUN-2000; 2000US-212210P.
PR 17-OCT-2000; 2000US-240816P.
PR 16-MAR-2001; 2001US-276248P.
PR 21-MAR-2001; 2001US-277379P.
PR 25-MAY-2001; 2001US-293499P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
PI WPI; 2002-114799/15.
DR
XX
XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for
PT the diagnosis and treatment of cancers and immune disorders -
PS
XX Claim 1; Page 2157-2158; 3148pp; English.
PS
XX This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BLyS) polypeptides. BLyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antineumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BLyS. The antibodies bind to BLyS
CC and so may be used to detect and quantitate the presence of BLyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BLyS. They may also be
CC administered to treat diseases associated with aberrant BLyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, Rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method
CC of the invention.
XX
XX
SQ Sequence 247 AA;
Query Match 92.1%; Score 82; DB 23; Length 247;
Best Local Similarity 88.2%; Pred. No. 9.8e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKQDSEKYYADSVRG 17
|||
50 NIKQDSEKYYADSVRG 66
Db
RESULT 15
ABP45583
ID ABP45583 standard; Protein; 251 AA.
XX
XX ABP45583;
AC
XX 19-AUG-2002 (first entry)
DT
XX
XX Human BLyS binding scFv SEQ ID 1594.
DE
XX
XX BLyS; B lymphocyte stimulator; TNF superfamily; human; cytosolic;
KM tumour necrosis factor; B cell proliferation; B cell differentiation;
KM immunosuppressive; immunostimulant; immunomodulatory; antineumatic;
KM antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KM systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KM common variable immunodeficiency; acquired immunodeficiency syndrome.

XX
XX Homo sapiens.
OS
XX MO200202641-A1.
PN
XX 10-JAN-2002.
PD
XX 15-JUN-2001; 2001WO-US19110.
PF
XX 16-JUN-2000; 2000US-212210P.
PR 17-OCT-2000; 2000US-240816P.
PR 16-MAR-2001; 2001US-276248P.
PR 21-MAR-2001; 2001US-277379P.
PR 25-MAY-2001; 2001US-293499P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
PI WPI; 2002-114799/15.
DR
XX
XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for
PT the diagnosis and treatment of cancers and immune disorders -
PS
XX Claim 1; Page 2305-2306; 3148pp; English.
PS
XX This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BLyS) polypeptides. BLyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antineumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BLyS. The antibodies bind to BLyS
CC and so may be used to detect and quantitate the presence of BLyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BLyS. They may also be
CC administered to treat diseases associated with aberrant BLyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method
CC of the invention.
XX
XX
SQ Sequence 251 AA;
Query Match 92.1%; Score 82; DB 23; Length 251;
Best Local Similarity 88.2%; Pred. No. 1e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKQDSEKYYADSVRG 17
|||
50 NIKQDSEKYYADSVRG 66
Db

Search completed: January 12, 2004, 06:58:52
Job time : 41.1719 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:56:19 ; Search time 11.5938 Seconds
(without alignments)
51.092 Million cell updates/sec

Title: US-09-829-495-63

Perfect score: 80

Sequence: 1 DKMEAYITPGADV 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	14	4	US-09-832-312-63
2	47	58.8	325	2	US-09-055-097-5
3	41	51.2	363	1	US-08-488-961-4
4	41	51.2	363	3	US-08-973-297-4
5	41	51.2	363	5	PCT-US96-06511-4
6	41	51.2	411	3	US-09-070-356-6
7	41	51.2	420	4	US-08-113-890-2
8	41	51.2	810	4	US-09-328-352-7181
9	40	50.0	117	1	US-07-942-245-32
10	40	50.0	117	1	US-07-942-245-36
11	40	50.0	117	1	US-07-942-245-38
12	39	48.8	310	4	US-09-252-991A-25177
13	39	48.8	546	4	US-09-252-991A-31673
14	39	48.8	596	1	US-08-565-386-11
15	39	48.8	1045	1	US-08-452-083-2
16	38	47.5	114	4	US-09-344-587-10
17	38	47.5	117	2	US-08-822-028-2
18	38	47.5	117	3	US-08-479-285-2
19	38	47.5	117	4	US-08-961-309-45
20	38	47.5	133	2	US-08-822-028-6
21	38	47.5	133	3	US-08-822-028-30
22	38	47.5	133	3	US-08-479-285-6
23	38	47.5	133	3	US-08-479-285-30
24	38	47.5	133	4	US-08-961-309-48
25	38	47.5	133	6	521996-17
26	38	47.5	134	2	US-08-822-028-10
27	38	47.5	134	3	US-08-479-285-10

28	38	47.5	134	4	US-08-961-309-49	Sequence 49, Appl
29	38	47.5	150	2	US-08-400-115-2	Sequence 2, Appl
30	38	47.5	271	2	US-08-400-115-4	Sequence 4, Appl
31	38	47.5	271	4	US-09-328-352-4737	Sequence 4, Appl
32	38	47.5	288	3	US-08-995-280C-2	Sequence 2, Appl
33	38	47.5	288	3	US-09-215-042-2	Sequence 2, Appl
34	38	47.5	495	4	US-09-252-991A-20739	Sequence 20739, A
35	38	47.5	7257	3	US-09-335-409-5	Sequence 5, Appl
36	38	47.5	7257	4	US-09-568-102-5	Sequence 5, Appl
37	38	47.5	7257	4	US-09-567-969-5	Sequence 5, Appl
38	38	47.5	7257	4	US-09-568-480-5	Sequence 5, Appl
39	38	47.5	7257	4	US-09-568-486-5	Sequence 5, Appl
40	38	47.5	7257	4	US-09-568-472-5	Sequence 5, Appl
41	38	47.5	7257	4	US-09-567-899-5	Sequence 5, Appl
42	37	46.2	363	4	US-08-488-961-7	Sequence 7, Appl
43	37	46.2	363	3	US-08-973-297-7	Sequence 7, Appl
44	37	46.2	363	5	PCT-US96-06511-7	Sequence 7, Appl
45	37	46.2	623	1	US-08-332-838-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-832-312-63
; Sequence 63, Application US/09832312
; Patent No. 6548741
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-63

Query Match
Best Local Similarity 100.0%; Score 80; DB 4; Length 14;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DKMEAYITPGADV 14
Db      1 DKMEAYITPGADV 14

RESULT 2
US-09-055-097-5
; Sequence 5, Application US/09050597
; Patent No. 5955282
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
```

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,097
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0490 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1150971
US-09-055-097-5

Query Match 58.8%; Score 47; DB 2; Length 325;
Best Local Similarity 58.3%; Pred. No. 4;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
DB 229 DRWPPYVTAGAF 240

RESULT 3
US-08-488-961-4
Sequence 4, Application US/08488961
Patent No. 5606042
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
TITLE OF INVENTION: Glycine and Phaseolus
TITLE OF INVENTION: alpha-D-Galactosidases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reising, Ethelington, Barnard & Perry
STREET: P.O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
ZIP: 48099-4390
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,961
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-320 (UMO)

TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-488-961-4

Query Match 51.2%; Score 41; DB 1; Length 363;
Best Local Similarity 46.2%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 13
DB 201 DKWASYAGPGWN 213

RESULT 4
US-08-973-297-4
Sequence 4, Application US/08973297
Patent No. 6184017
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
TITLE OF INVENTION: Glycine and Phaseolus
TITLE OF INVENTION: alpha-D-Galactosidases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6184017thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,297
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994,00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5050
TELEFAX: (810) 539-5055
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-973-297-4

Query Match 51.2%; Score 41; DB 3; Length 363;
Best Local Similarity 46.2%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 13
DB 201 DKWASYAGPGWN 213

RESULT 5
PCT-US96-06511-4
Sequence 4, Application PC/TUS9606511

GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
APPLICANT: Walker, John C.
TITLE OF INVENTION: Glycine and Phaseolus
TITLE OF INVENTION: alpha-D-Galactosidases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 Northwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06511
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994.00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5050
TELEFAX: (810) 539-5055
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US96-06511-4

Query Match 51.2%; Score 41; DB 5; Length 363;
Best Local Similarity 46.2%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAPD 13
Db 201 DKWASYAGPGWN 213

RESULT 6
US-09-070-356-6
Sequence 6, Application US/09070356
Patent No. 6228631
GENERAL INFORMATION:
APPLICANT: Alex Zhu
APPLICANT: Jack Goldstein
TITLE OF INVENTION: Recombinant a-N-
TITLE OF INVENTION: Acetylglucosaminidase
TITLE OF INVENTION: Enzyme and cDNA Encoding
TITLE OF INVENTION: Said Enzyme
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rochstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,356
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/037,248
FILING DATE: March 26, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 63475/12
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: cDNA to mRNA
HYPOTHETICAL: no
ANTI-SENSE: yes
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: guar plant *Cyamopsis tetragonoloba*
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: library
POSITION IN GENOME: unknown
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: guar a-galactosidase
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Overbeek et al
TITLE: Cloning and Nucleotide Sequence of
Patent No. 6228631
TITLE: The a-Galactosidase cDNA From
TITLE: *Cyamopsis tetragonoloba* (guar)
JOURNAL: Plant Molecular Biology
VOLUME: 13
PAGES: 541-550
DATE: 1989
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-070-356-6

Query Match 51.2%; Score 41; DB 3; Length 411;
Best Local Similarity 46.2%; Pred. No. 48;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAPD 13
Db 248 DKWASYAGPGWN 260

RESULT 7
US-08-113-890-2
Sequence 2, Application US/08113890
Patent No. 6329191
GENERAL INFORMATION:

APPLICANT: IVY, JOHN M.
APPLICANT: CLEMENTS, DAVID E.
TITLE OF INVENTION: RECOMBINANT COFFEE BEAN
TITLE OF INVENTION: ALPHA-GALACTOSIDASE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/113,890
FILING DATE: 30-AUG-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 4733-0002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-113-890-2

Query Match 51.2%; Score 41; DB 4; Length 420;
Best Local Similarity 46.2%; Pred. No. 49;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFD 13
DB 258 DKWASYAGPGGMN 270

RESULT 8
US-09-328-352-7181
Sequence 7181, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Bretton et al.
TITLE OF INVENTION: NOCUEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7181
LENGTH: 810
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7181

Query Match 51.2%; Score 41; DB 4; Length 810;
Best Local Similarity 77.8%; Pred. No. 95;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 AYITPGAFD 13
DB 160 SYATPGAFD 168

RESULT 9

US-07-942-245-32
Sequence 32, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrie, Mion, Zimm, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-32

Query Match 50.0%; Score 40; DB 1; Length 117;
Best Local Similarity 46.2%; Pred. No. 20;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
DB 46 EWEAYITSGSFTI 58

RESULT 10
US-07-942-245-36
Sequence 36, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrie, Mion, Zimm, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 workstation
OPERATING SYSTEM: UNIX

SOFTWARE: In house
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/942,245
 FILING DATE: 09-SEP-1992
 CLASSIFICATION: 530
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 293-7060
 TELEFAX: (202) 293-7860
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 36:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 117 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-07-942-245-36

Query Match 50.0%; Score 40; DB 1; Length 117;
 Best Local Similarity 46.2%; Pred. No. 20;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
 DB 46 EWWAYISGSFTI 58

RESULT 11
 US-07-942-245-38
 Sequence 38, Application US/07942245
 Patent No. 5639641
 GENERAL INFORMATION:
 APPLICANT: PEDERSEN, Jan T.
 APPLICANT: SEARLE, Stephen M.J.
 APPLICANT: REES, Anthony R.
 APPLICANT: ROGUSKA, Michael A.
 APPLICANT: GUILD, Braydon C.
 TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
 TITLE OF INVENTION: ANTIBODIES
 NUMBER OF SEQUENCES: 522
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
 STREET: 2100 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States
 ZIP: 20037-3202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: HP 9000/700 Workstation
 OPERATING SYSTEM: UNIX
 SOFTWARE: In house
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/942,245
 FILING DATE: 09-SEP-1992
 CLASSIFICATION: 530
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 293-7060
 TELEFAX: (202) 293-7860
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 117 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-07-942-245-38

Query Match 50.0%; Score 40; DB 1; Length 117;
 Best Local Similarity 46.2%; Pred. No. 20;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
 DB 46 EWWAYISGSFTI 58

DB 46 EWWAYISGSFTI 58

RESULT 12
 US-09-252-991A-25177
 Sequence 25177, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 25177
 LENGTH: 310
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-25177

Query Match 48.8%; Score 39; DB 4; Length 310;
 Best Local Similarity 53.8%; Pred. No. 76;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 DKWEAYITPGADV 13
 DB 275 DAWASWPGAVE 287

RESULT 13
 US-09-252-991A-31673
 Sequence 31673, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 31673
 LENGTH: 546
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-31673

Query Match 48.8%; Score 39; DB 4; Length 546;
 Best Local Similarity 50.0%; Pred. No. 13e+02;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGADV 14
 DB 529 DRGAYDKGAYDI 542

RESULT 14
 US-08-565-386-11
 Sequence 11, Application US/08565386
 Patent No. 5741697
 GENERAL INFORMATION:
 APPLICANT: Bayoil, Patrik M.
 APPLICANT: Hsiao, Ru-ching
 TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI

Query Match 50.0%; Score 40; DB 1; Length 117;
 Best Local Similarity 46.2%; Pred. No. 20;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
 DB 46 EWWAYISGSFTI 58

RESULT 14
 US-08-565-386-11
 Sequence 11, Application US/08565386
 Patent No. 5741697
 GENERAL INFORMATION:
 APPLICANT: Bayoil, Patrik M.
 APPLICANT: Hsiao, Ru-ching
 TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI

Query Match 50.0%; Score 40; DB 1; Length 117;
 Best Local Similarity 46.2%; Pred. No. 20;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
 DB 46 EWWAYISGSFTI 58

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Roches
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/565,386
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 176/60040
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-565-386-11

Query Match 48.8%; Score 39; DB 1; Length 596;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEAYITPG 10
|:|:|:|:
Db 94 DNWESFITGG 103

RESULT 15
US-08-452-083-2
Sequence 2, Application US/08452083
Patent No. 5756327
GENERAL INFORMATION:
APPLICANT: Sasseanfar, Mandana
TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL ISOLEUCYL-tRNA
TITLE OF INVENTION: SYNTHETASE GENES, TESTER STRAINS AND ASSAYS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,083
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/305,765
FILING DATE: 13-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: C194-08B
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1045 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-452-083-2

Query Match 48.8%; Score 39; DB 1; Length 1045;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKWEAYIT 8
|:|:|:|:
Db 144 DEMQAYVT 151

Search completed: January 12, 2004, 07:03:59
Job time : 12.5938 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2004, 06:56:19 ; Search time 14.0781 Seconds
(without alignments)
51.092 Million cell updates/sec

Title: US-09-829-495-62

Perfect score: 89

Sequence: 1 NIKODGSEKXYADSVRG 17

Scoring table: BLOSUM62

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	4	US-09-832-312-62
2	82	92.1	98	1	US-07-942-245-37
3	82	92.1	117	3	US-08-545-809A-95
4	82	92.1	312	4	US-09-079-029-10
5	69	77.5	116	3	US-08-545-809A-134
6	67	75.3	117	3	US-09-240-274-141
7	65	73.0	114	3	US-08-545-809A-124
8	65	73.0	126	1	US-08-476-349A-95
9	65	73.0	98	1	US-08-211-202-118
10	64	71.9	116	1	US-08-211-202-141
11	64	71.9	117	3	US-08-545-809A-115
12	64	71.9	119	3	US-08-331-398A-46
13	64	71.9	119	2	US-08-331-397B-46
14	64	71.9	119	2	US-08-759-804A-46
15	64	71.9	119	3	US-09-227-693-46
16	64	71.9	120	1	US-08-211-202-135
17	64	71.9	120	1	US-07-942-245-35
18	64	71.9	123	3	US-08-983-607-38
19	64	71.9	123	1	US-08-276-852-72
20	64	71.9	125	1	US-08-276-852-73
21	64	71.9	125	1	US-08-276-852-74
22	64	71.9	125	1	US-08-276-852-75
23	64	71.9	125	1	US-08-276-852-76
24	64	71.9	125	1	US-08-276-852-77
25	64	71.9	125	1	US-08-276-852-77
26	64	71.9	125	1	US-08-899-575-72
27	64	71.9	125	1	US-08-899-575-73

28	64	71.9	125	1	US-08-899-575-74	Sequence 74, Appl
29	64	71.9	125	1	US-08-899-575-75	Sequence 75, Appl
30	64	71.9	125	1	US-08-899-575-76	Sequence 76, Appl
31	64	71.9	125	1	US-08-899-575-77	Sequence 77, Appl
32	64	71.9	125	1	US-08-899-575-72	Sequence 72, Appl
33	64	71.9	125	1	US-08-899-575-73	Sequence 73, Appl
34	64	71.9	125	1	US-08-899-575-74	Sequence 74, Appl
35	64	71.9	125	1	US-08-899-575-75	Sequence 75, Appl
36	64	71.9	125	1	US-08-899-575-76	Sequence 76, Appl
37	64	71.9	125	1	US-08-899-575-77	Sequence 77, Appl
38	64	71.9	125	5	PCT-US95-08743-72	Sequence 72, Appl
39	64	71.9	125	5	PCT-US95-08743-73	Sequence 73, Appl
40	64	71.9	125	5	PCT-US95-08743-74	Sequence 74, Appl
41	64	71.9	125	5	PCT-US95-08743-75	Sequence 75, Appl
42	64	71.9	125	5	PCT-US95-08743-76	Sequence 76, Appl
43	64	71.9	125	5	PCT-US95-08743-77	Sequence 77, Appl
44	64	71.9	128	1	US-08-476-039-96	Sequence 96, Appl
45	64	71.9	128	1	US-08-476-349A-96	Sequence 96, Appl

ALIGNMENTS

```
RESULT 1
US-09-832-312-62
; Sequence 62, Application US/09832312
; Patent No. 6548741
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-62
Query Match      100.0%  Score 89; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 NIKODGSEKXYADSVRG 17
Db      1 NIKODGSEKXYADSVRG 17
RESULT 2
US-07-942-245-37
; Sequence 37, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
```

STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-37

Query Match 92.1%; Score 82; DB 1; Length 98;
Best Local Similarity 88.2%; Pred. No. 4.9e-07;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYYVDSVVG 17
DB 50 NIKDGSSEKYYVDSVVG 66

RESULT 3
US-08-545-809A-95
Sequence 95, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Taauku
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-95

Query Match 92.1%; Score 82; DB 3; Length 117;
Best Local Similarity 88.2%; Pred. No. 6e-07;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYYVDSVVG 17
DB 69 NIKDGSSEKYYVDSVVG 85

RESULT 4
US-09-079-029-10
Sequence 10, Application US/09079029
Patent No. 6342369
GENERAL INFORMATION:
APPLICANT: Adams, Camilla W.
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Chundharapai, Anan
TITLE OF INVENTION: Apo-2 Receptor
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatlin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,029
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P1101R2
TELEPHONE: 650/225-5416
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-079-029-10

Query Match 92.1%; Score 82; DB 4; Length 312;
Best Local Similarity 88.2%; Pred. No. 1.8e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYYVDSVVG 17
DB 89 NIKDGSSEKYYVDSVVG 105

RESULT 5
US-08-545-809A-134
Sequence 134, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Taauku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE

TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-134

Query Match 77.5%; Score 69; DB 3; Length 116;
Best Local Similarity 76.5%; Pred. No. 0.00011;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYVADSVRG 17
:|||||:|||||:
Db 68 DIKDGSEKYYVDSVKG 84

RESULT 6
US-09-240-274-141
; Sequence 141, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: RH(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 141
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) antibody clone SH17
US-09-240-274-141

Query Match 75.3%; Score 67; DB 3; Length 117;
Best Local Similarity 76.5%; Pred. No. 0.00024;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYVADSVRG 17
:|||||:|||||:
Db 50 NIKDGSSEKYYVDSVKG 66

RESULT 7
US-08-545-809A-124
; Sequence 124, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Matsuda, Fumihiko
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,809A
; FILING DATE: 27-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP93/00603
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 06501/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 124:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-545-809A-124

Query Match 73.0%; Score 65; DB 3; Length 114;
Best Local Similarity 70.6%; Pred. No. 0.00053;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NIKDGSSEKYVADSVRG 17
:|||||:|||||:
Db 68 HIMDGSSEKYYVDSVKG 84

RESULT 8
US-08-478-039-95
; Sequence 95, Application US/08478039
; Patent No. 5681722
; GENERAL INFORMATION:
; APPLICANT: Newman, Roland A.
; APPLICANT: Hanna, Nabil
; APPLICANT: Raab, Ronald W.
; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATTHIS
; STREET: 699 Prince St.
; CITY: Alexandria

STATE: VA
COUNTRY: USA
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,039
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,072
FILING DATE: 25-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/912,292
FILING DATE: 10-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/856,281
FILING DATE: 23-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,064
FILING DATE: 25-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teskin Esq., Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-160
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
POSITION IN GENOME:
CHROMOSOME/SEGMENT: RF SJ1
US-08-478-039-95

Query Match 73.0%; Score 65; DB 1; Length 126;
Best Local Similarity 75.0%; Pred. No. 0.00059;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 IKODGSEKYYADSVRG 17
Db 51 ISDDGSNKYYADSVKG 66

RESULT 9
US-08-476-349A-95
Sequence 95, Application US/08476349A
Patent No. 5750105
GENERAL INFORMATION:
APPLICANT: Newman, Roland A.
APPLICANT: Hanna, Nabill
APPLICANT: Raab, Ronald W.
TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince St.
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,349A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,072
FILING DATE: 25-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/912,292
FILING DATE: 10-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/856,281
FILING DATE: 23-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,064
FILING DATE: 25-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teskin Esq., Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-161
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
POSITION IN GENOME:
CHROMOSOME/SEGMENT: RF SJ1
US-08-476-349A-95

Query Match 73.0%; Score 65; DB 1; Length 126;
Best Local Similarity 75.0%; Pred. No. 0.00059;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 IKODGSEKYYADSVRG 17
Db 51 ISDDGSNKYYADSVKG 66

RESULT 10
US-08-211-202-118
Sequence 118, Application US/08211202
Patent No. 5565332
GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, Hendicus Reneus Jacobus Mattheus
APPLICANT: BAIER, Michael
APPLICANT: JESPEERS, Laurent Stephane Anne Therese
APPLICANT: WINTER, Gregory Paul
TITLE OF INVENTION: Production of chimeric antibodies - a
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESSES:
ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

      1 FILING DATE: 25-SEP-1991
      2 PRIOR APPLICATION DATA:
      3 APPLICATION NUMBER: GB 9206318.9
      4 FILING DATE: 24-MAR-1992
      5 PRIOR APPLICATION DATA:
      6 APPLICATION NUMBER: GB 9206372.6
      7 FILING DATE: 24-MAR-1992
      8 PRIOR APPLICATION DATA:
      9 APPLICATION NUMBER: PCT/GB92/00883
     10 FILING DATE: 15-MAY-1992
     11 ATTORNEY/AGENT INFORMATION:
     12 NAME: David W. Clough
     13 REGISTRATION NUMBER: 36,107
     14 REFERENCE/DOCKET NUMBER: 28111/31960
     15 TELECOMMUNICATION INFORMATION:
     16 TELEPHONE: 312-474-6300
     17 TELEFAX: 312-474-0448
     18 TELEX: 25-3856
     19 INFORMATION FOR SEQ ID NO: 141:
     20 SEQUENCE CHARACTERISTICS:
     21 LENGTH: 116 amino acids
     22 TYPE: amino acid
     23 TOPOLOGY: linear
     24 MOLECULE TYPE: protein
     25 US-08-211-202-141

Query Match 71.9%; Score 64; DB 1; Length 116;
Best Local Similarity 75.0%; Pred. No. 0.0008;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Cy 2 IKODSEKYYADSVRG 17
   | ||| |||||:|
Db 51 ISYDGSNKYYADSVKG 66

RESULT 12
US-08-545-809A-115
Sequence 115, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:

```

LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-115

Query Match 71.9%; Score 64; DB 3; Length 117;
Best Local Similarity 75.0%; Pred. No. 0.00081;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 IKODGSEKYYADSVRG 17
DB 70 ISYDGSNKYYADSVKG 85

RESULT 13
US-08-331-398A-46
Sequence 46, Application US/08331398A
Patent No. 5608039
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Williamson, Mark
APPLICANT: Fitzgerald, David
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
TITLE OF INVENTION: and their uses (as amended)
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,398A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..119
OTHER INFORMATION:
OTHER INFORMATION: 56P1'CL Variable Heavy chain (V-H)"
US-08-331-398A-46
Query Match 71.9%; Score 64; DB 1; Length 119;
Best Local Similarity 75.0%; Pred. No. 0.00082;

Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 IKODGSEKYYADSVRG 17
DB 51 ISYDGSNKYYADSVKG 66

RESULT 14
US-08-331-397B-46
Sequence 46, Application US/08331397B
Patent No. 5981726
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Benhar, Itai
TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-Specific Antibody Fragments, Fusion Proteins, and Uses Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,397B
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126120US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..119
OTHER INFORMATION:
OTHER INFORMATION: 56P1'CL Variable Heavy chain (V-H)"
US-08-331-397B-46
Query Match 71.9%; Score 64; DB 2; Length 119;
Best Local Similarity 75.0%; Pred. No. 0.00082;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 IKODGSEKYYADSVRG 17
DB 51 ISYDGSNKYYADSVKG 66
RESULT 15
US-08-759-804A-46
Sequence 46, Application US/08759804A


```

: Patent No. 5390296
: GENERAL INFORMATION:
: APPLICANT: Pastan, Ira
: APPLICANT: Willingham, Mark
: APPLICANT: Fitzgerald, David J.
: APPLICANT: Brinkmann, Ulrich
: APPLICANT: Pai, Lee
: TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
: TITLE OF INVENTION: Fusion Proteins, and Uses Thereof
: NUMBER OF SEQUENCES: 68
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/759,804A
: FILING DATE: 03-DEC-1996
: CLASSIFICATION: 53C
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/331,398
: FILING DATE: 28-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/767,331
: FILING DATE: 30-SEP-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/596,289
: FILING DATE: 12-OCT-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Weber, Ellen L.
: REGISTRATION NUMBER: 32,762
: REFERENCE/DOCKET NUMBER: 015280-126140US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 46:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 119 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..119
: OTHER INFORMATION: /note="Human fetal immunoglobulin
: OTHER INFORMATION: 56Pl CL Variable Heavy chain (V-H)"
: US-08-759-804A-46
:
: Query Match 71.9% Score 64; DB 2; length 119;
: Best Local Similarity 75.0%; Pred. No. 0.00082;
: Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0
:
: QY 2 IKDGSSEKRYADSVRG 17
: ||| |||||:|
: Db 51 ISYDSNKKYADSVKVG 66

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:49 ; Search time 12.6875 Seconds
(without alignments)
106.117 Million cell updates/sec

Title: US-09-829-495-63
Perfect score: 80
Sequence: 1 DKWEAYITGAFDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	57.5	383	2 B95252	alcohol dehydrogen
2	46	57.5	383	2 H98116	probable alcohol d
3	46	57.5	484	2 T38463	probable aminotran
4	45	56.2	765	2 S76795	hypothetical prote
5	44	55.0	175	2 T00186	dUTP diaphosphatase
6	44	55.0	682	2 T50840	U4/U6-associated R
7	44	55.0	683	2 T50839	U4/U6 small nuclea
8	44	55.0	867	1 C64785	outer membrane ush
9	44	55.0	869	2 B90703	probable outer mem
10	44	55.0	869	2 B85553	hypothetical prote
11	43	53.8	653	2 F85620	partial fimbrial u
12	43	53.8	653	2 H90756	partial fimbrial u
13	43	53.8	866	1 C64834	probable outer mem
14	42	52.5	108	2 PH1015	ig heavy chain V r
15	42	52.5	238	2 E90142	conserved hypochet
16	42	52.5	317	2 F87634	LysR-family transc
17	41	51.2	293	2 AF0475	rifin PFB055C - m
18	41	51.2	312	2 E71624	alpha-galactosidas
19	41	51.2	378	2 T50781	phoH-related prote
20	41	51.2	380	2 A75328	alpha-galactosidas
21	41	51.2	411	2 S07472	alpha-galactosidas
22	41	51.2	422	2 T06388	alpha-galactosidas
23	41	51.2	604	2 AB1989	hypothetical prote
24	41	51.2	684	2 H82296	c-di-GMP phosphodi
25	41	51.2	796	2 T23238	hypothetical prote
26	40	50.0	721	2 C71014	hypothetical prote
27	40	50.0	1444	2 AD1602	DNA polymerase III
28	40	50.0	1444	2 AH1239	DNA polymerase III
29	40	50.0	3158	2 T17483	peptide synthetase

30	39	48.8	115	2 AF1247	hypothetical prote
31	39	48.8	115	2 AB1610	hypothetical prote
32	39	48.8	148	2 F71207	hypothetical prote
33	39	48.8	179	2 D71957	hypothetical prote
34	39	48.8	264	2 F70828	hypothetical prote
35	39	48.8	366	2 B87446	hypothetical prote
36	39	48.8	367	2 E91182	probable outer mem
37	39	48.8	367	2 A86029	probable fimbrial
38	39	48.8	434	2 T47748	alpha-galactosidas
39	39	48.8	495	2 G83325	probable sulfatase
40	39	48.8	592	2 PS0197	phospholipase C (E
41	39	48.8	596	1 U00345	major capsid prote
42	39	48.8	766	2 T01900	isp4 protein homol
43	39	48.8	998	2 T04842	protein kinase AK4
44	39	48.8	1041	2 E70760	probable 11es prot
45	38	47.5	71	2 T43545	DNA-directed RNA p

ALIGNMENTS

RESULT 1
B95252 alcohol dehydrogenase, iron-containing [imported] - Streptococcus pneumoniae (strain TIC
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C/Accession: B95252
R/Retelin, H.; Nelson, K.E.; Paulsen, I.T.; Bisen, J.A.; Read, T.D.; Peterson, S.; Heid
son, J.D.; Umayam, L.A.; White, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A./Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A/Reference number: A95000; PMID:21357209; PMID:11463916
A/Accession: B95252
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-383 <KUR>
A/Cross-references: GB:AE005672; PIDN:AAK76211.1; PID:g14973668; GSPDB:GN00164; TIGR:SP4
A/Experimental source: strain TIGR4
C/Genetics:
A/Genetic: SP2157
C/Superfamily: lactaldehyde reductase; lactaldehyde reductase homology
Query Match 57.5%; Score 46; DB 2; Length 383;
Best Local Similarity 72.7%; Pred. No. 3.7; 1; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 4 EAYITGAFDV 14
199 EALVTPGAYDV 209
RESULT 2
H98116 probable alcohol dehydrogenase (EC 1.1.1.1) [imported] - Streptococcus pneumoniae (strain
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C/Accession: H98116
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; W
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain K6.
A/Reference number: A97872; PMID:21429245; PMID:11544234
A/Accession: H98116
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-383 <KUR>
A/Cross-references: GB:AE007317; PIDN:AAU00765.1; PID:g15459663; GSPDB:GN00174
C/Genetics:
A/Genetic: adh2

C/Superfamily: lactaldehyde reductase; lactaldehyde reductase homology
C/Keywords: oxidoreductase; NAD

Query Match 57.5%; Score 46; DB 2; Length 383;
Best Local Similarity 72.7%; Pred. No. 3.7;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAYITPGAFDV 14
||:||||:|
199 EALVITPGADV 209

RESULT 3

T38463
probable aminotransferase (EC 2.6.1.-) - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C/Accession: T38463

R/Harris, D.; McDonald, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996

A/Reference number: Z21794

A/Accession: T38463

A/Status: preliminary; translated from GB/EMBL/DBDJB

A/Molecule type: DNA

A/Residues: 1-484 <HAR>

A/Cross-references: EMBL:Z69368; PIDN:CAA93294.1; GSPDB:GN00066; SPDB:SPAC27F1.05C

A/Experimental source: strain 972h-; cosmid c27F1

A/Genetic: SPDB:SPAC27F1.05C

A/Map position: 1

C/Superfamily: beta-alanine-pyruvate transaminase

C/Keywords: aminotransferase

Query Match 57.5%; Score 46; DB 2; Length 484;
Best Local Similarity 50.0%; Pred. No. 4.8;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFDV 14
||:||||:|
204 EKWKOTSPAPFDV 217

RESULT 4

S76795
hypothetical protein - Synechocystis sp. (strain PCC 6803)

C/Species: Synechocystis sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C/Accession: S76795

R/Kaneho, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamiyu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

S.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S76795

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-765 <KAN>

A/Cross-references: EMBL:D90916; GB:AB001339; NID:G1653715; PIDN:BA18707.1; PID:dl01944

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 56.2%; Score 45; DB 2; Length 765;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEAYITPG 11
||:||||:|
160 DGMWITPG 170

RESULT 5

T00186

dmp diphosphatase (EC 3.6.1.23) - Staphylococcus aureus phage phi PVL

N/Alternate names: dmpase

C/Species: Staphylococcus aureus phage phi PVL

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002

C/Accession: T00186

R/Kaneho, U.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kamio, Y.

Biosci. Biotechnol. Biochem. 61, 1960-1962, 1997

A/Title: Pantone-Valentine leukocidin genes in a phage-like particle isolated from mitomy

A/Reference number: Z14119; MUID:98067870; PMID:9404084

A/Accession: T00186

A/Status: translated from GB/EMBL/DBDJB

A/Molecule type: DNA

A/Residues: 1-175 <KAN>

A/Cross-references: EMBL:AB009866; NID:G3341907; PIDN:BA31926.1; PID:G3341960

C/Function:

A/Description: catalyzes the hydrolysis of dmp to dmp

C/Superfamily: retroviral proteolase

C/Keywords: hydrolase; nucleotide metabolism

Query Match 55.0%; Score 44; DB 2; Length 175;
Best Local Similarity 77.8%; Pred. No. 3.5;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 YITPGAFDV 14
|||||:
108 YITPGAFDI 116

RESULT 6

T50840
U4/U6-associated RNA splicing factor [imported] - human

C/Species: Homo sapiens (man)

C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000

C/Accession: T50840

R/Wang, A.; Forman-Kay, J.; Luo, Y.; Luo, M.; Chow, Y.H.; Plumb, J.; Friesen, J.D.; Tsu

Hum. Mol. Genet. 6, 2117-2126, 1997

A/Title: Identification and characterization of human genes encoding Hprp3 and Hprp4p,

A/Reference number: Z12232

A/Accession: T50840

A/Status: preliminary; translated from GB/EMBL/DBDJB

A/Molecule type: mRNA

A/Residues: 1-682 <MAN>

A/Cross-references: EMBL:AF001947; PIDN:AAC09069.1

C/Genetics:

A/Genetic: PRP3

Query Match 55.0%; Score 44; DB 2; Length 682;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
||:||||:|
390 WDSYITPGFDL 401

RESULT 7

T50839

U4/U6 small nuclear ribonucleoprotein hPrp3 [imported] - human

C/Species: Homo sapiens (man)

C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000

C/Accession: T50839

R/Horowitz, D.S.; Kobayashi, R.; Krainer, A.R.

RNA 3, 1374-1387, 1997

A/Title: A new cyclophilin and the human homologues of yeast Prp3 and Prp4 form a compl

A/Reference number: Z12233

A/Accession: T50839

A/Status: preliminary; translated from GB/EMBL/DBDJB

A/Molecule type: mRNA

A/Residues: 1-683 <HOR>

A/Cross-references: EMBL:AF016370; PIDN:AAC51926.1

C/Genetics:

A/Genetic: hPrp3

Query Match 55.0%; Score 44; DB 2; Length 683;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
:::|||||
Db 390 WDSYIIFNGFDL 401

RESULT 8
C64785
Outer membrane usher protein sfmd precursor - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C:Accession: C64785
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
S:Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; PMID:97426617; PMID:9278503
A:Accession: C64785
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-867 <BLAT>
A:Cross-References: GB:AE000159; GB:U00096; NID:g1786739; PIDN:AGC73634.1; PID:g1786744;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: smfd
C:Superfamily: outer membrane usher protein fimd
C:Keywords: membrane protein

Query Match 55.0%; Score 44; DB 1; Length 867;
Best Local Similarity 41.7%; Pred. No. 20;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
:::|||||
Db 309 YQSYVSPGAFEI 320

RESULT 9
B90703
Probable outer membrane protein EC80594 [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: B90703
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: B90703
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-869 <HAY>
A:Cross-References: GB:BA000007; PIDN:BAR34017.1; PID:g13360052; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC80594
C:Superfamily: outer membrane usher protein fimd

Query Match 55.0%; Score 44; DB 2; Length 869;
Best Local Similarity 41.7%; Pred. No. 20;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
:::|||||
Db 309 YQSYVSPGAFEI 320

RESULT 10
B85553
hypothetical protein sfmd [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85553
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lm, A.; Dialante, E.; Potamousts, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; PMID:21074935; PMID:11206551
A:Accession: B85553
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-869 <STO>
A:Cross-References: GB:AE005174; NID:g12513432; PIDN:AGS54889.1; GSPDB:GN00145; UWGP:206
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: sfmd
C:Superfamily: outer membrane usher protein fimd

Query Match 55.0%; Score 44; DB 2; Length 869;
Best Local Similarity 41.7%; Pred. No. 20;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
:::|||||
Db 309 YQSYVSPGAFEI 320

RESULT 11
F85620
Partial fimbrial usher protein [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: F85620
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lm, A.; Dialante, E.; Potamousts, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; PMID:21074935; PMID:11206551
A:Accession: F85620
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-653 <STO>
A:Cross-References: GB:AE005174; NID:g12513419; PIDN:AGS5426.1; GSPDB:GN00145; UWGP:212
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1289

Query Match 53.8%; Score 43; DB 2; Length 653;
Best Local Similarity 41.7%; Pred. No. 22;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
:::|||||
Db 95 YQSYVSPGAFEI 106

RESULT 12
H90756
Partial fimbrial usher protein [similarity] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C:Accession: H90756
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: H90756
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-653 <HAY>
A:Cross-References: GB:BA000007; PIDN:BAR34447.1; PID:g13360483; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:
A:Gene: ECs1024

Query Match 53.8%; Score 43; DB 2; Length 653;
Best Local Similarity 41.7%; Pred. No. 22;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 3 WEAYITPGAPDV 14
DB 95 YQTVSPGAFEI 106

RESULT 13

C64834 probable outer membrane usher protein ycbS - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002

C/Accession: C64834

R/Battner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A: Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; MUID:97426617; PMID:9278503

A/Accession: C64834

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-866 <BLAT>

A/Cross-references: GB:AE000196; GB:U00096; NID:gl187169; PIND:AACT4026.1; PID:gl187172;

A/Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: ycbS

C/Superfamily: outer membrane usher protein fimb

C/Keywords: ATP; fimbria; membrane protein; nucleotide binding; P-loop; transport protei

F/15-31/Domain: transmembrane #status predicted <TMM>

F/642-649/Region: nucleotide-binding motif A (P-loop)

Query March 53.8%; Score 43; DB 1; Length 866;
Best Local Similarity 41.7%; Pred. No. 30;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 3 WEAYITPGAPDV 14
DB 308 YQTVSPGAFEI 319

RESULT 14

PH1015

IG heavy chain V region (clone 111.55) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C/Accession: PH1015

R/Tillman, D.W.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A/Title: Both IGM and IGC anti-DNA antibodies are the products of clonally selective B c

A/Reference number: PH0971; MUID:92381444; PMID:1512540

A/Accession: PH1015

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-108 <TIL>

A/Experimental source: B cell, strain [NZB x NZW]F1

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F/15-98/Domain: immunoglobulin homology <IMM>

Query Match 52.5%; Score 42; DB 2; Length 108;
Best Local Similarity 46.2%; Pred. No. 4.6;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 KWEAYITPGAPDV 14
DB 46 EWVAYISSGFNI 58

RESULT 15

E90142

conserved hypothetical protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C/Accession: E90142

R/she, Q.J. Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, E

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: E90142

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-238 <KUR>

A/Cross-references: GB:AE006641; NID:gl3813160; PIND:AAK40396.1; GSPDB:GN00155

C:Genetics:

A:Gene: SSO0031

Query Match 52.5%; Score 42; DB 2; Length 238;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 KWEAYITPGAPD 13
DB 53 RYVAYITPGAFE 64

Search completed: January 12, 2004, 07:03:01
Job time: 16.6875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2004, 06:47:49 ; Search time 7.65625 seconds
(without alignments)

85.992 Million cell updates/sec

Title: US-09-829-495-63

Perfect score: 80
Sequence: 1 DKMEAYITPGAFDV 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	57.5	484	YAV5_SCHPO	Q10174 schizosacch
2	44	55.0	867	SPMD_ECOLI	P77468 escherichia
3	43	53.8	333	DCUP_THEAC	Q9h189 thermoplasma
4	43	53.8	866	YC8S_ECOLI	P75857 escherichia
5	41	51.2	293	HDPF_YERPE	Q82327 yerinia pe
6	41	51.2	378	AGAL_COPAR	Q42556 coffea arab
7	41	51.2	411	AGAL_CYATE	P14749 cyamopsis t
8	41	51.2	796	YRS7_CAEEL	Q10003 caenorhabd
9	40	50.0	870	FIMD_SALTY	Q37924 salmonella
10	40	50.0	1444	DPO3_LISIN	Q92634 listeria in
11	40	50.0	1444	DPO3_LISMO	Q87391 listeria mo
12	39	48.8	283	PHLC_BACCE	P09598 bacillus ce
13	39	48.8	283	PHLD_BACCE	P33376 bacillus ce
14	39	48.8	366	Q9A7Y2_CAUOBACT	Q9A7Y2 caulobacter
15	39	48.8	595	VPI_BPCHP	P19192 bacterioph
16	39	48.8	1041	SVI_MYCTU	Q10765 mycobacteri
17	38	47.5	66	RPOH_SULTO	Q95794 sulfolobus
18	38	47.5	71	RPEX_SCHPO	O13877 schizosacch
19	38	47.5	109	Y030_MYCTU	P71600 mycobacteri
20	38	47.5	208	RS3A_SULSO	Q9u4d4 sulfolobus
21	38	47.5	258	SSRA_ARATH	P45334 arabidopsis
22	38	47.5	422	PYRC_HALNT	Q9hnb9 halobacteri
23	38	47.5	756	MALC_SCHPO	Q10768 schizosacch
24	37.5	46.9	599	RROF_PAMV	P29154 pea enation
25	37	46.2	66	RPOH_SULSO	Q98028 sulfolobus
26	37	46.2	70	RPEX_SCHPO	P22133 saccharomyc
27	37	46.2	225	MTBD_METKA	Q32864 methanopyru
28	37	46.2	329	XYNA_THEAU	P23360 thermosacch
29	37	46.2	341	Q9EA_CLOPE	Q8x115 clostridium
30	37	46.2	463	Y030_NPYAC	P41334 autographa
31	37	46.2	555	FTPI_BOMMO	P49667 bombyx mori
32	37	46.2	568	PHAC_CHRVO	Q92n12 chromobacte
33	37	46.2	574	TYRO_PODAN	Q92396 podospora a

34	37	46.2	632	1	GAAT_HUMAN	Q9un88 homo sapien
35	37	46.2	702	1	GLGB_HUMAN	Q04446 homo sapien
36	37	46.2	811	1	RPOD_NEUTN	P33541 neurospora
37	37	46.2	831	1	RPOF_GELSP	O03685 gelasinospo
38	37	46.2	1001	1	RPCR_MOUSE	O970x5 mus musculu
39	37	46.2	2254	1	CCAG_RAT	O54898 rattus norv
40	37	46.2	2377	1	CCAG_HUMAN	O43497 homo sapien
41	37	46.2	4289	1	TENX_HUMAN	P22105 homo sapien
42	36	45.0	138	1	SCHB_STRAA	Q05362 streptomyce
43	36	45.0	209	1	UPP_STRAA	O99686 staphylococ
44	36	45.0	303	1	SPAO_SALTI	Q56022 salmonella
45	36	45.0	361	1	Q9EA_RHITO	Q98m55 rhizobium 1

ALIGNMENTS

RESULT 1
YAV5_SCHPO STANDARD; PRT; 484 AA.
AC Q10174;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable aminotransferase C27FL.05c (EC 2.6.1.-).
GN SPAC27FL.05c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grynolprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goftau A., Cadieu E., Dreano S., Gloux S., Leleure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forstberg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Petachkin J.,
RA Shpakovskii G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -I- COFACTOR: Pyridoxal phosphate (Potential).
CC -I- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent
amino transferases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; Z69368; CAA93294.1; --
 DR PIR; T38463; T38463.
 DR HSSP; P12995; 1013.
 DR GeneDB; SPAC271.05C; --
 DR InterPro; IPR005814; Aminoctrans_3.
 DR Pfam; PF00202; aminotran_3; 1.
 DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
 DR Hypothetical protein; Transferase; Aminoctransferase;
 KM Pyridoxal phosphate.
 FT BINDING 305 305 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 484 AA; 53190 MW; 5DB9A13BA33C497 CRC64;
 QY Query Match 57.5%; Score 46; DB 1; Length 484;
 Best Local Similarity 50.0%; Pred. No. 3;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Db 204 EKWQKYQSPAFDV 217
 QY 1 DKWEAYITPGAFDV 14
 ID DCUP_THEAC STANDARD; PRT; 867 AA.
 AC P77468; P77133;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Outer membrane usher protein sfmd precursor.
 GN SFMD OR B0532.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=362;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden W.A., Rose D.D.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474 (1997).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
 RA Federapfel N., Hyman R., Kaiman S., Komp C., Kurdi O., Lew H., Lin D.,
 RA Nathath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RT "Sequence of minutes 4-25 of Escherichia coli.";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF SPINA FIMBRIAL
 CC SUBUNIT: ACROSS THE OUTER MEMBRANE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
 CC (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE000159; AAC73634.1; --
 DR EMBL; U82598; AAB40730.1; --
 DR EMBL; U82664; AAB40285.1; --
 DR PIR; C64785; C64785.
 DR EcoGene; EGI3883; sfmd.
 DR InterPro; IPR000015; Fimb_usher.
 DR Pfam; PF00577; Usher; 1.
 DR PROSITE; PS01151; FIMBRIAL_USHER; 1.

KM Outer membrane; Transmembrane, Fimbria; Transport; Signal;
 KW Complete proteome.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 867 OUTER MEMBRANE USHER PROTEIN SFMD.
 FT DISULFID 840 862 POTENTIAL.
 SQ SEQUENCE 867 AA; 95677 MW; DF8591D0BEC4205A CRC64;
 QY Query Match 55.0%; Score 44; DB 1; Length 867;
 Best Local Similarity 41.7%; Pred. No. 11;
 Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
 Db 309 YQSYVSPGAFEI 320
 QY 3 WEAYITPGAFDV 14
 ID DCUP_THEAC STANDARD; PRT; 333 AA.
 AC Q9HLB9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
 GN HEME OR TA0310.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma
 RT acidophilum.";
 RL Nature 407:508-513 (2000).
 CC -1- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4
 CC CO(2).
 CC
 CC -1- PATHWAY: Porphyrin biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL445063; CAC11455.1; --
 DR HSSP; P06132; URO.
 DR HAMAP; MF 00218; -; 1.
 DR InterPro; IPR006361; Heme.
 DR InterPro; IPR000257; Uro_decarboxyls.
 DR Pfam; PF01208; URO-D; 1.
 DR ProDom; PD003225; Uro_decarboxyls; 1.
 DR TIGRFAMs; TIGR01464; heme; 1.
 DR PROSITE; PS00906; UROD_1; 1.
 DR PROSITE; PS00907; UROD_2; FALSE NEG.
 KW lase; Decarboxylase; Porphyrin biosynthesis; Complete proteome.
 SQ SEQUENCE 333 AA; 37785 MW; 3DBE510768C0FF5B CRC64;
 QY Query Match 53.8%; Score 43; DB 1; Length 333;
 Best Local Similarity 38.5%; Pred. No. 6; 8;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Db 199 DSWAGYISPEYE 211
 QY 1 DKWEAYITPGAFDV 13
 ID DCUP_THEAC STANDARD; PRT; 333 AA.
 AC Q9HLB9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
 GN HEME OR TA0310.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma
 RT acidophilum.";
 RL Nature 407:508-513 (2000).
 CC -1- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4
 CC CO(2).
 CC
 CC -1- PATHWAY: Porphyrin biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL445063; CAC11455.1; --
 DR HSSP; P06132; URO.
 DR HAMAP; MF 00218; -; 1.
 DR InterPro; IPR006361; Heme.
 DR InterPro; IPR000257; Uro_decarboxyls.
 DR Pfam; PF01208; URO-D; 1.
 DR ProDom; PD003225; Uro_decarboxyls; 1.
 DR TIGRFAMs; TIGR01464; heme; 1.
 DR PROSITE; PS00906; UROD_1; 1.
 DR PROSITE; PS00907; UROD_2; FALSE NEG.
 KW lase; Decarboxylase; Porphyrin biosynthesis; Complete proteome.
 SQ SEQUENCE 333 AA; 37785 MW; 3DBE510768C0FF5B CRC64;


```
RESULT 4
YCBS_ECOLI STANDARD; PRT; 866 AA.
ID YCBS_ECOLI
AC P75857.1
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical outer membrane usher protein ycbS precursor.
GN YCBS OR B0940.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905237;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitegawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sasaki G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horinouchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- FUNCTION: COULD BE INVOLVED IN THE EXPORT AND ASSEMBLY OF THE
CC PUTATIVE YCB0 FIMBRIAL SUBUNIT ACROSS THE OUTER MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE000196; AAC74026.1; -.
DR EMBL, D90732; BAA35695.1; -.
DR PIR, C64834; C64834.
DR EcoGene, EGI3711; ycbS.
DR InterPro, IPR000015; Fimb_usher.
DR Pfam, PF00577; Usher.1.
DR PROSITE, PS01151; FIMBRIAL_USHER.1.
KW Hypothetical protein; Outer membrane; Transmembrane; Fimbria;
KW Transport; Signal; Complete proteome.
FT SIGNAL 1 35
FT CHAIN 36 866 HYPOTHETICAL OUTER MEMBRANE USHER PROTEIN
FT FT YCBS.
FT SEQUENCE 866 AA; 95241 MW; 0004DC5E91F5796 CRC64;
SQ
Query Match 53.8%; Score 43; DB 1; Length 866;
Best Local Similarity 41.7%; Pred. No. 17;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
```

```
RESULT 5
HDFR_YERPE STANDARD; PRT; 293 AA.
ID HDFR_YERPE
AC O82AA7.1
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcriptional regulator hdfs (H-NS-dependent flhDC regulator).
GN HDFR OR YPO3904 OR Y0332.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Tlball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatina M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holtroyd S., Jagers K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skellon J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Retherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
CC -1- FUNCTION: Negatively regulates the transcription of the flagellar
CC master operon flhDC by binding to the upstream region of the
CC operator (By similarity).
CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AJ14159; CAC93370.1; -.
DR EMBL, AE013633; AAM83923.1; -.
DR PIR, AF0475; AF0475.
DR HAMAP, MF_01233; -; 1.
DR InterPro, IPR000647; HTH_LysR.
DR InterPro, IPR005119; LysR_subst.
DR Pfam, PF00126; HTH_1.1.
DR Pfam, PF03466; LysR_substrate; 1.
DR PRINTS, PR00039; HTHLYSR.
DR PROSITE, PS00044; HTH_LYSR_FAMILY.1.
KW Transcription regulation; Repressor; DNA-binding; Complete proteome.
FT DNA_BIND 18 37
FT FT H-T-H MOTIF (BY SIMILARITY).
FT SEQUENCE 293 AA; 33932 MW; F5EC224D0D7E71EB CRC64;
SQ
Query Match 51.2%; Score 41; DB 1; Length 293;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 6
ID AGAL_COPAR STANDARD; PRT; 378 AA.
AC Q42656;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alpha-galactosidase precursor (EC 3.2.1.22) (Melibiase) (Alpha-D-galactoside galactohydrolyase).
OS Coffea arabica (Coffee).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamids; Gentianales; Rubiaceae; Ichoridaceae; Coffeaceae; Coffea.
OC NCBI_TaxID=13443;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-34; 215-231 AND 373-378.
RC TISSUE=Seed;
RX MEDLINE=94193002; PubMed=8144030;
RA Zhu A., Goldstein J.;
RT "Cloning and functional expression of a cDNA encoding coffee bean alpha-galactosidase."
RT Gene 140:227-231(1994).
RL
CC -1- FUNCTION: PREFERENTIALLY CLEAVES ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDE LINKAGES. INVOLVED IN THE HYDROLYSIS OF THE GALACTOMANNAN. IT SPLITTS ALPHA-LINKED GALACTOSE MOIETIES. IT IS PARTICULARLY SUITABLE FOR THE HYDROLYSIS OF GUAR GUM TO A GUM WITH IMPROVED GELLING PROPERTIES. CAN CLEAVE TERMINAL ALPHA-1,3-LINKED GALACTOSE RESIDUES RESPONSIBLE FOR BLOOD GROUP B SPECIFICITY FROM THE SURFACE OF ERYTHROCYTES THEREBY CONVERTING THESE CELLS SEROLOGICALLY TO GROUP O.
CC -1- CATALYTIC ACTIVITY: Melibiase + H(2)O = galactose + glucose.
CC -1- SIMILARITY: BELONGS TO FAMILY 27 OF GLYCOSYL HYDROLASES.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: L27992; AAA3022.1; -
CC PIR: T50781; T50781.
DR InterPro: IPR002241; Glyco_hydro_27.
DR InterPro: IPR000111; Glyco_hydro_GHD.
DR Pfam: PF02065; Melibiase; 1.
DR PRINTS: PR00740; GLHYDRLSE27.
DR ProDom: PD002572; Glyco_hydro_GHD; 1.
DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
KM Hydrolyase; Glycosidase; Signal.
FT SIGNAL 1 15
FT CHAIN 16 378 ALPHA-GALACTOSIDASE.
FT ACT SITE 287 287 POTENTIAL.
FT SEQUENCE 378 AA; 41310 MM; 9ECT610BFD60A3 CRC64;
SQ
Query Match 51.2%; Score 41; DB 1; Length 378;
Best Local Similarity 46.2%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFD 13
DB 216 DKWASYAGPGWN 228

RESULT 7
ID AGAL_CYATE STANDARD; PRT; 411 AA.
AC P14749;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Alpha-galactosidase precursor (EC 3.2.1.22) (Melibiase) (Alpha-D-galactoside galactohydrolyase). (Cluster bean).
OS Cyamopsis tetragonoloba (Guar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Indigotereae;
OC Cyamopsis.
OC NCBI_TaxID=3832;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aleurone;
RX MEDLINE=91370836; PubMed=2577496;
RA Overbeek N., Fellinger A.J., Toonen M.Y., van Wassenar D., Verrips C.T.;
RT "Cloning and nucleotide sequence of the alpha-galactosidase cDNA from Cyamopsis tetragonoloba (guar)";
RT Plant Mol. Biol. 13:541-550(1989).
RL
RN [2]
RP SEQUENCE OF 48-57 AND 172-178.
RC TISSUE=Seed;
RX Hughes S.G., Overbeek N., Robinson S., Pollock K., Smeets F.L.M.;
RT "Messenger RNA from isolated aleurone cells directs the synthesis of an alpha-galactosidase found in the endosperm during germination of guar (Cyamopsis tetragonoloba) seed."
RT Plant Mol. Biol. 11:783-789(1988).
RL
CC -1- FUNCTION: INVOLVED IN THE HYDROLYSIS OF THE GALACTOMANNAN. IT SPLITTS ALPHA-LINKED GALACTOSE MOIETIES. IT IS PARTICULARLY SUITABLE FOR THE HYDROLYSIS OF GUAR GUM TO A GUM WITH IMPROVED GELLING PROPERTIES. PREFERENTIALLY CLEAVES ALPHA-1,6 GLYCOSIDE LINKAGES.
CC -1- CATALYTIC ACTIVITY: Melibiase + H(2)O = galactose + glucose.
CC -1- SIMILARITY: BELONGS TO FAMILY 27 OF GLYCOSYL HYDROLASES.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: X14619; CAA2772.1; -
CC PIR: S07472; S07472.
DR InterPro: IPR002241; Glyco_hydro_27.
DR InterPro: IPR000111; Glyco_hydro_GHD.
DR Pfam: PF02065; Melibiase; 1.
DR PRINTS: PR00740; GLHYDRLSE27.
DR ProDom: PD002572; Glyco_hydro_GHD; 1.
DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
KM Hydrolyase; Glycosidase; Signal; Glycoprotein.
FT SIGNAL 1 24
FT PROPEP 25 47
FT CHAIN 48 411
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT ACT SITE 319 319 POTENTIAL.
FT SEQUENCE 411 AA; 45135 MM; 5B1715858DAB1E CRC64;
SQ
Query Match 51.2%; Score 41; DB 1; Length 411;
Best Local Similarity 46.2%; Pred. No. 18;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFD 13
DB 248 DKWASYAGPGWN 260

RESULT 8
ID YRS7_CAEEL STANDARD; PRT; 796 AA.
AC Q10003;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Hypothetical 30.8 kDa protein T05H10.7 in chromosome II.

GN T05H10.7.

OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Pelodierinae; *Caenorhabditis*.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2.

RA Ligniting J., Thomas K.;

RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; 247811; CAA87788.1; -

DR EMBL; 247812; CAA87788.1; JOINED.

DR EMBL; 247812; CAA87796.1; -

DR EMBL; 247811; CAA87796.1; JOINED.

DR PIR; T23238; T23238.

DR WormPep; T05H10.7; CE03637.

DR InterPro; IPR004129; GDPD.

DR Pfam; PF03009; GDPD; 1.

KW Hypothetical protein.

FT DOMAIN 38 42 POLY-GLU.

FT DOMAIN 67 70 POLY-GLU.

FT DOMAIN 524 527 POLY-GLU.

FT SEQUENCE 796 AA; 90831 MW; 7BDF8E0A4D2AA9F1 CRC64;

SQ

Query Match

Best Local Similarity 51.2%; Score 41; DB 1; Length 796;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DKWEAVITP 9

DB 163 DKWEAVLHP 171

RESULT 9

FIND SALTY STANDARD; PRT; 870 AA.

AC P37924;

DT 01-OCT-1994 (Rel. 30, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Outer membrane usher protein fimb precursor.

GN FIMD OR STM0546.

OS *Salmonella typhimurium*.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; *Salmonella*.

OX NCBI_TaxID=602;

RN [1]

RP SEQUENCE FROM N.A.

RC Swenson D.L., Clegg S.;

RA "The complete nucleotide sequence of a *Salmonella typhimurium* type 1

RA fimbrial gene cluster."

RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

RX MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Lattelle P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,

RA Leonard S., Nguyen K., Scott K., Holmes A., Grewal N., Mulvaney E.,

RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium

RT LT2."

RL Nature 413:852-856(2001).

CC -1- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF FIMA FIMBRIAL

CC SUBUNIT: ACROSS THE OUTER MEMBRANE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane

CC (By similarity).

CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; L19338; AAA75419.1; -

DR EMBL; AE008721; AAL19500.1; -

DR StGene; SG10505; fimb.

DR InterPro; IPR000015; Fimb_usher.

DR Pfam; PF00577; Usher; 1.

DR PROSITE; PS01151; FIMBRIAL_USHER; 1.

KW Outer membrane; Transmembrane; Fimbria; Transport; Signal;

KW Complete proteome.

FT SIGNAL 1 27 POTENTIAL.

FT CHAIN 28 870 OUTER MEMBRANE USHER PROTEIN FIMD.

FT DISULFID 843 865 POTENTIAL.

FT CONFLICT 265 265 V -> F (IN REF. 1).

FT CONFLICT 361 361 G -> A (IN REF. 1).

FT SEQUENCE 870 AA; 95127 MW; A6B000139F654A29 CRC64;

SQ

Query Match

Best Local Similarity 50.0%; Score 40; DB 1; Length 870;

Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 3 WEAVITPGAFV 14

DB 310 YQSVSPGAFV 321

RESULT 10

DPO3 LISIN STANDARD; PRT; 1444 AA.

AC Q92C34;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE DNA polymerase III polC-type (EC 2.7.7.7) (PolIII).

GN POLC OR LIN1357.

OS *Listeria innocua*.

OC Bacteria; Firmicutes; Bacillales; *Listeriaceae*; *Listeria*.

OX NCBI_TaxID=1642;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CLIP 11262 / Serovar 6a.

RX MEDLINE=21537279; PubMed=11679669;

RA Glaeser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,

RA Entian K.-D., Fehli H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunz F., Kurapkat G.,

RA Madeno E., Maitournam A., Mara Vicente J., Ng E., Nedjari H.,

RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

RA "Comparative genomics of *Listeria species*."

RL Science 294:849-852(2001).

CC -1- FUNCTION: Required for replicative DNA synthesis. This DNA

CC polymerase also exhibits 3' to 5' exonuclease activity (By

```

CC similarity).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. POLC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL596168; CAC96588.1; -.
CC PIR: AD1602; AD1602.
CC Listliet: LIN01357; -.
CC HAMAP: MF_00356; -.
CC InterPro: IPR006054; DnaQ.
CC InterPro: IPR006055; Exonuclease.
CC InterPro: IPR003141; PNP_C.
CC InterPro: IPR006308; PolC_gram_pos.
CC InterPro: IPR004365; trna_anti.
CC Pfam: PF00929; Exonuclease; 1.
CC Pfam: PF02811; PNP_C; 1.
CC Pfam: PF02231; PNP_N; 1.
CC Pfam: PF01336; trna_anti; 1.
CC SMART: SM00479; EXOIII; 1.
CC SMART: SM00481; POLI1AC; 1.
CC TIGRFAMs: TIGR00573; dnaq; 1.
CC TIGRFAMs: TIGR01405; polC_gram_pos; 1.
CC Transferase: DNA-directed DNA polymerase; DNA replication; Hydrolase;
CC Nuclease; Exonuclease; Complete proteome.
CC DOMAIN 428 592 EXONUCLEASE.
CC SQ SEQUENCE 1444 AA; 162904 MW; D4B78BCA39D9AD95 CRC64;

Query Match 50.0%; Score 40; DB 1; Length 1444;
Best Local Similarity 70.0%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DKMEAYITPG 10
Db 460 DKFEAFIDPG 469

RESULT 11
DPO3 LISMO STANDARD; PRT; 1444 AA.
AC 08Y7G1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase III polC-type (EC 2.7.7.7) (PolIII).
GN POLC OR IM01320.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21517279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Checouani F., Couve E., de Darvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duseurget O.,
RA Ertian K.-D., Fathi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstlek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlutener T., Simoes N., Tierrez A.,

```

```

RA Vazquez-Boland J.-A., Voets H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
CC -1- FUNCTION: Required for replicative DNA synthesis. This DNA
CC polymerase also exhibits 3' to 5' exonuclease activity (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. POLC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL591978; CAC9398.1; -.
CC PIR: AH1239; AH1239.
CC Listliet: LM001320; -.
CC HAMAP: MF_00356; -.
CC InterPro: IPR006054; DnaQ.
CC InterPro: IPR006055; Exonuclease.
CC InterPro: IPR004013; PNP_C.
CC InterPro: IPR003141; PNP_N.
CC InterPro: IPR006308; PolC_gram_pos.
CC InterPro: IPR004365; trna_anti.
CC Pfam: PF00929; Exonuclease; 1.
CC Pfam: PF02811; PNP_C; 1.
CC Pfam: PF02231; PNP_N; 1.
CC Pfam: PF01336; trna_anti; 1.
CC SMART: SM00479; EXOIII; 1.
CC SMART: SM00481; POLI1AC; 1.
CC TIGRFAMs: TIGR00573; dnaq; 1.
CC TIGRFAMs: TIGR01405; polC_gram_pos; 1.
CC Transferase: DNA-directed DNA polymerase; DNA replication; Hydrolase;
CC Nuclease; Exonuclease; Complete proteome.
CC DOMAIN 428 592 EXONUCLEASE.
CC SQ SEQUENCE 1444 AA; 162732 MW; DBB3F57FBDP3815B CRC64;

Query Match 50.0%; Score 40; DB 1; Length 1444;
Best Local Similarity 70.0%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DKMEAYITPG 10
Db 460 DKFEAFIDPG 469

RESULT 12
PHLC_BACCE STANDARD; PRT; 283 AA.
AC P05398;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine
DE cholinephosphohydrolase) (Cereolysin A).
GN PLC.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=SE-1;
RX MEDLINE=6833678; PubMed=3137122;
RA Johansen T., Holm T., Gudgal P.H., Sletten K., Haugli F.B., Little C.;
RT "Cloning and sequencing of the gene encoding the phosphatidylcholine-
RT preferring phospholipase C of Bacillus cereus.";
RL Gene 65:293-304 (1988).

```

RN [2] SEQUENCE FROM N.A.
 RP STRAIN-VKM B-164;
 RC MEDLINE=93249510; PubMed=8387306;
 RA Kuzmin N.P., Gavrilenko I.V., Krutov V.M., Karpov A.V.;
 RT "Nucleotide sequence of phospholipase C and sphingomyelinase genes
 from *Bacillus cereus* BKM-B164 (letter).";
 RL Bioorg. Khim. 19:133-138(1993).
 RN [3]
 RP SEQUENCE OF 166-283 FROM N.A.
 RC STRAIN-1AM 1208;
 RX MEDLINE=88296483; PubMed=2841128;
 RA Yamada A., Tsukagoshi N., Uda S., Sasaki T., Makino S., Nakamura S.,
 RA Little C., Tomita M., Ikezawa H.;
 RT "Nucleotide sequence and expression in *Escherichia coli* of the gene
 coding for sphingomyelinase of *Bacillus cereus*.";
 RL Eur. J. Biochem. 175:213-220(1988).
 RN [4]
 RP SEQUENCE OF 39-65.
 RX MEDLINE=78043154; PubMed=72664;
 RA Otnaess A.-B., Little C., Sletten K., Wallin R., Johnsen S.,
 RA Fløgersrud R., Prydz H.;
 RT "Some characteristics of phospholipase C from *Bacillus cereus*.";
 RL Eur. J. Biochem. 79:459-468(1977).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
 RX MEDLINE=89159438; PubMed=2493587;
 RA Hough E., Hansen L.K., Birkenes B., Jynge K., Hansen S., Hordvik A.,
 RA Little C., Dodson E., Derewenda Z.;
 RT "High-resolution (1.5 Å) crystal structure of phospholipase C from
Bacillus cereus.";
 RL Nature 338:357-360(1989).
 CC -1- FUNCTION: REQUIRED, WITH SPHINGOMYELINASE TO EFFECT TARGET CELL
 LYSIS (HEMOLYSIS).
 CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
 diacylglycerol + choline phosphate.
 CC -1- COFACTOR: Binds 3 zinc ions.
 CC -1- SUBUNIT: Monomer.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X64141; CAA45502.1; -;
 CC EMBL: X12854; CAA3332.1; -;
 CC EMBL: X12711; CAA31213.1; -;
 CC EMBL: X64140; CAA45501.1; ALT_TERM.
 CC PIR: S18978; PS0197.
 CC PDB: 1AH7; 10-DEC-97.
 CC InterPro: IP0001531; Zn dep. PLPC.
 CC Pfam: PF00882; Zn dep. PLPC.
 CC PRINTS: PR00479; FRPHEPLPASEC.
 CC ProDom: PD003946; Zn dep. PLPC; 1.
 CC PROSITE: PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
 CC KMW HydroLase; Zinc; Signal; Zymogen; Hemolysis; 3D-structure.
 CC FT SIGNAL 1 24 POTENTIAL.
 CC FT PROPEP 1 38
 CC FT CHAIN 35 283 PHOSPHOLIPASE C.
 CC FT METAL 39 39 ZINC 3.
 CC FT METAL 52 52 ZINC 3.
 CC FT METAL 93 93 ZINC 1.
 CC FT METAL 107 107 ZINC 1.
 CC FT METAL 156 156 ZINC 1.
 CC FT METAL 160 160 ZINC 1 AND 3.
 CC FT METAL 166 166 ZINC 2.
 CC FT METAL 180 180 ZINC 2.
 CC FT METAL 184 184 ZINC 2.

FT	VARIANT	212	212	E -> D (IN STRAIN 1AM1208).
FT	VARIANT	226	226	S -> A (IN STRAIN 1AM1208).
FT	VARIANT	239	239	K -> R (IN STRAIN 1AM1208).
FT	VARIANT	282	282	D -> N (IN STRAIN 1AM1208).
FT	HELIX	47	49	
FT	HELIX	51	65	
FT	HELIX	72	80	
FT	HELIX	81	81	
FT	HELIX	82	91	
FT	TURN	92	94	
FT	TURN	96	103	
FT	HELIX	105	107	
FT	STRAND	110	110	
FT	STRAND	111	114	
FT	STRAND	115	115	
FT	TURN	119	120	
FT	HELIX	124	140	
FT	TURN	141	142	
FT	HELIX	144	161	
FT	TURN	162	162	
FT	TURN	164	169	
FT	TURN	172	173	
FT	TURN	177	178	
FT	HELIX	179	190	
FT	HELIX	191	194	
FT	TURN	203	203	
FT	HELIX	210	223	
FT	TURN	224	224	
FT	HELIX	225	227	
FT	TURN	228	228	
FT	HELIX	231	240	
FT	TURN	241	242	
FT	HELIX	244	280	
SO	SEQUENCE	283 AA;	32383 MW;	ACS452EF2E22B19 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 283;
 Best Local Similarity 66.7%; Pred. No. 27;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEAYITP 9
 DB 248 DKWEAYITP 256

RESULT 13
 PHLD_BACCE STANDARD; PRT; 283 AA.
 ID PHLD_BACCE
 AC P33376;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine
 de cholinephosphodiolase) (Cereolysin A).
 GN CERA.
 OS *Bacillus cereus*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_Taxid=1396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GP-4;
 RX MEDLINE=89123149; PubMed=2536680;
 RA Gilmore M.S., Cruz-Rodriguez A.L., Leimister-Waechter M., Kreft J.,
 RA Goebel W.;
 RT "A *Bacillus cereus* cytolytic determinant, cereolysin AB, which
 comprises the phospholipase C and sphingomyelinase genes: nucleotide
 sequence and genetic linkage.";
 RL J. Bacteriol. 171:744-753(1989).
 RN [2]
 RP SEQUENCE OF 50-142 FROM N.A.
 RA Gilmore M.S., Gilmore K.S., Goebel W.;
 RT "A new strategy for ordered DNA sequencing based on a novel method for
 the rapid purification of near-milligram quantities of a cloned
 restriction fragment.";

```

RL Gene Anal.Tech. 2:108-114(1995)
CC -1- FUNCTION: REQUIRED, WITH SPHINGOMYELINASE TO EFFECT TARGET CELL
CC LYSIS (HEMOLYSIS).
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC -1- COFACTOR: Binds 3 zinc ions.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24149; AAA91819.1; -.
DR EMBL; M35411; AAA2524.1; -.
DR PIR; S18978; PS0197.
DR HSSP; P09598; IAH7.
DR InterPro; IPR001531; Zn_dep_PLPC.
DR Pfam; PF00882; Zn_dep_PLPC; 1.
DR PRINTS; PR00479; BRPBPPLPASEC.
DR ProDom; PS003946; Zn_dep_PLPC; 1.
DR PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
KW Hydrolyase; Zinc; Signal; Zymogen; Hemolysis.
FT SIGNAL 1 24
FT PROPEP 25 38 BY SIMILARITY.
FT CHAIN 39 283 PHOSPHOLIPASE C.
FT METAL 39 39 ZINC 3.
FT METAL 52 52 ZINC 3.
FT METAL 93 93 ZINC 1.
FT METAL 107 107 ZINC 1.
FT METAL 156 156 ZINC 1.
FT METAL 160 160 ZINC 1 AND 3.
FT METAL 166 166 ZINC 2.
FT METAL 180 180 ZINC 2.
FT METAL 184 184 ZINC 2.
SQ SEQUENCE 283 AA; 32373 MW; 24D2E753402A6A44 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 283;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKMEAYITP 9
Db 248 DKMEAYITP 256

RESULT 14
QUEA_CAUCR
ID QUEA_CAUCR STANDARD; PRT; 366 AA.
AC Q9A7Y2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S-adenosylmethionine: tRNA ribosyltransferase-isoenzyme (EC 5.5.1.1)
DE (Queuosine biosynthesis protein queA).
DE QUEA OR CC1587.
GN QUEA
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka J., Nelson W.C., Newton A., Stephens C., Phade N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,

```

CC	-1- PATHWAY: Quenosine biosynthesis.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC	-1- SIMILARITY: BELONGS TO THE OUEA FAMILY.

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outpost -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).

DR	EMBL; AE005833; AAC23566.1; .
DR	PIR; B87446; B87446.
DR	TIGR; CC1587; .
DR	HAMAP: MF_00113; -. 1.
DR	InterPro: IPR003699; Quenosine synth.
DR	Pfam; PF02547; Quenosine_synth_1.
DR	TIGRFamS; TIGR00113; quea; 1.
KW	Quenosine biosynthesis; Transferase; Isomerase; Complete proteome.
SQ	SEQUENCE 366 AA; 39817 MW; A40DBEE37E70C7C3 CRC64;
Cy	Query Match 48.8%; Score 39; DB 1; Length 366;
Dc	Best Local Similarity 50.0%; Pred. NO. 35;
Matches	5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
1	DKWEAVITPG 10 : : : : : :
Db	97 DRMSAFMRFG 106
RESULT 15	
ID_VPL_BPCHP	STANDARD; PRT; 595 AA.
AC_P19I92:	
DT_01-NOV-1990	(Rel. 16, Created)
DT_01-NOV-1990	(Rel. 16, Last sequence update)
DT_28-FEB-2003	(Rel. 41, Last annotation update)
DE_Protein Vpl	(ORF).
OS_Bacteriophage Chp1.	
OC_Viruses; ssDNA viruses; Microviridae; Microvirus.	
OX_NCBI_TaxID=12367;	
RN [1]	
RP_SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.	
RA_MEDLINE=90111716; PubMed=2607241;	
RA_Storey C.C.; Lueher M.; Richmond S.J.;	
RT "Analysis of the complete nucleotide sequence of Chp1, a phage which	
infects Avian Chlamydia psittaci.";	
J. Gen. Virol. 70:3381-3390(1989).	
RL J. Gen. Virol. 70:3381-3390(1989).	

CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL outpost -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC or send an email to license@isb-sib.ch).	

DR_EMBL; D00624; BAA00515.1; .	
DR_InterPro; IPR003514; Caspid_F.	
DR_Pfam; PF02305; Pfam_F_1.	
KW_Structural protein.	
FT_INIT MET 0	
SEQUENCE 595 AA; 66850 MW; AE71D8024EA967ED CRC64;	

Query Match	48.8%	Score 39	DB 1	Length 595
Best Local Similarity	60.0%	Pred. No. 55		
Matches	6	Conservative	2	Mismatches 2
				Indels 0
				Gaps 0

QY	1	DKWEAYITPG	10
Db	93	DNWESFITGG	102

Search completed: January 12, 2004, 06:59:39
Job time : 10.6562 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:54:04 ; Search time 28.4375 Seconds
(without alignments)
127.041 Million cell updates/sec

Title: US-09-829-495-63

Perfect score: 80
Sequence: 1 DKMEAYITPGAFDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	65.0	167	16 Q98D81	Q98D81 rhizobium 1
2	47	58.8	244	5 Q8S2S9	Q8S2S9 drosophila
3	47	58.8	325	5 Q9W4N0	Q9W4N0 drosophila
4	47	58.8	325	5 Q24157	Q24157 drosophila
5	46	57.5	383	16 Q97N98	Q97N98 streptococc
6	46	57.5	383	16 Q8DN25	Q8DN25 streptococc
7	45	56.2	765	16 P74599	P74599 synectocyst
8	44	55.0	175	9 Q8SDV3	Q8SDV3 bacterioph
9	44	55.0	175	9 Q80D91	Q80D91 staphylococ
10	44	55.0	354	5 Q9VCT5	Q9VCT5 drosophila
11	44	55.0	669	2 Q70012	Q70012 streptomyce
12	44	55.0	669	16 Q9KX04	Q9KX04 streptomyce
13	44	55.0	682	4 Q43395	Q43395 homo sapien
14	44	55.0	683	4 Q43446	Q43446 homo sapien
15	44	55.0	683	11 Q9D6C6	Q9D6C6 mus musculu
16	44	55.0	683	11 Q922U1	Q922U1 mus musculu

17	44	55.0	869	16 Q8XCT4	Q8XCT4 escherichia
18	43	53.8	653	16 Q8XDD1	Q8XDD1 escherichia
19	43	53.8	799	10 Q94GFB	Q94GFB oryza sativ
20	43	53.8	1087	10 Q8SAX4	Q8SAX4 oryza sativ
21	42	52.5	80	2 Q8KLB0	Q8KLB0 rhizobium e
22	42	52.5	238	17 Q981B6	Q981B6 sulfobolus
23	42	52.5	243	16 Q992R9	Q992R9 streptococc
24	42	52.5	317	16 Q9A3U0	Q9A3U0 caulobacter
25	42	52.5	705	17 Q8U121	Q8U121 pyrococcus
26	42	52.5	873	2 Q8GC99	Q8GC99 citrobacter
27	42	52.5	1522	10 Q942K3	Q942K3 oryza sativ
28	41	51.2	245	16 Q8P0C8	Q8P0C8 streptococc
29	41	51.2	245	16 Q8E6F3	Q8E6F3 streptococc
30	41	51.2	245	16 Q8E0T5	Q8E0T5 streptococc
31	41	51.2	295	17 Q8U371	Q8U371 pyrococcus
32	41	51.2	312	5 Q96116	Q96116 plasmidum
33	41	51.2	329	16 Q8K717	Q8K717 streptococc
34	41	51.2	380	16 Q9K3Y1	Q9K3Y1 deinococcus
35	41	51.2	383	11 Q9CUL6	Q9CUL6 mus musculu
36	41	51.2	395	10 Q8S2Q4	Q8S2Q4 oryza sativ
37	41	51.2	422	10 Q39811	Q39811 glycine max
38	41	51.2	446	16 Q9RIU8	Q9RIU8 streptomyce
39	41	51.2	604	16 Q8WY11	Q8WY11 anabaena sp
40	41	51.2	684	16 Q9KU70	Q9KU70 vibrio chol
41	41	51.2	696	5 Q955T9	Q955T9 drosophila
42	41	51.2	741	10 Q9FDJ2	Q9FDJ2 arabidopsis
43	41	51.2	948	6 Q9RTD7	Q9RTD7 trichosurus
44	41	51.2	2030	5 Q9VWN8	Q9VWN8 drosophila
45	40	50.0	97	10 Q947G0	Q947G0 glycine max

ALIGNMENTS

RESULT 1

ID	Q98D81	PRELIMINARY;	PRT;	167 AA.
AC	Q98D81;			
DT	01-OCT-2001 (TREMBLrel. 18, Created)			
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Transcription regulator.			
ML	ML14820.			
OS	Rhizobium loti (Mesorhizobium loti).			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Phyllobacteriaceae; Mesorhizobium.			
OX	NCBI_TaxId=381;			
RP	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MAFF303099;			
RX	MEDLINE=21082930; PubMed=11214968;			
RA	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,			
RA	Watanabe A., Idekawa K., Ishikawa A., Kawashima K., Kimura T.,			
RA	Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,			
RA	Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,			
RA	Takeuchi C., Yamada M., Tabeta S.;			
RT	"Complete genome structure of the nitrogen-fixing symbiotic bacterium			
RT	Mesorhizobium loti."			
RL	DNA Res. 7:331-338 (2000).			
DR	EMBL; AP003005; BAB51390.1; -			
DR	EMBL; IPR002579; DUF25.			
DR	InterPro; IPR006311; Tac.			
DR	Pfam; PF01641; Seir1.			
DR	ProDom; PD004057; DUF25; 1.			
DR	TIGRFAms; TIGR01409; TAT_signal_seq; 1.			
DR	TIGRFAms; TIGR00357; TIGR00357; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 167 AA; 17876 MW; 1E0F7327BE2646CC CRC64;			

Query Match 65.0%; Score 52; DB 16; Length 167;
Best Local Similarity 57.1%; Pred. No. 0.39;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKWEAYITPGAFDV 14
 Db 45 DKWALISPAAFDV 58

RESULT 2

Q85ZS9 PRELIMINARY; PRT; 244 AA.
 AC Q85ZS9;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE ID10479P.
 GN BRN OR EG:EG0007.6 OR CG4934.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Fafan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY070536; AAL48007.1; -.
 DR FlyBase; FBgn0000221; brn.
 DR InterPro; IPR002659; Glyco_trans_31.
 DR Pfam; PF01762; Galactosyl_T; 1.
 SQ SEQUENCE 244 AA; 28448 MW; FF78A628F12C3C6B CRC64;

Query Match 58.8%; Score 47; DB 5; Length 244;
 Best Local Similarity 58.3%; Pred. No. 4.4;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DKWEAYITPGAF 12
 Db 148 DRWPPYVTAGAF 159

RESULT 3

Q9W4NO PRELIMINARY; PRT; 325 AA.
 AC Q9W4NO;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE BRN protein.
 GN BRN OR EG:EG0007.6 OR CG4934.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Peiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brodtier P.,
 RA Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dong L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matesi B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazotto M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AEO03430; AAP45918.1; -.
 DR FlyBase; FBgn0000221; brn.
 DR InterPro; IPR002659; Glyco_trans_31.
 DR Pfam; PF01762; Galactosyl_T; 1.
 SQ SEQUENCE 325 AA; 37619 MW; 0DF89B720F43657B CRC64;

Query Match 58.8%; Score 47; DB 5; Length 325;
 Best Local Similarity 58.3%; Pred. No. 6.1;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DKWEAYITPGAF 12
 Db 229 DRWPPYVTAGAF 240

RESULT 4

Q24157 PRELIMINARY; PRT; 325 AA.
 AC Q24157;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Neurogenic secreted signaling protein precursor (Brainiac protein).
 GN BRN OR EG:EG0007.6.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R;
 RA MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Bartell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
 RA Dreano S., Gloux S., Leleure V., Mottier S., Galibert P., Borkova D.,
 RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,

RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablo B.,
 RA Modolell J., Peter A., Schoetler P., Werner M., Mourikoti F.,
 RA Beinert N., Dove G., Schaefer U., Jaekle H., Bicheton A.,
 RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
 RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;
 RT "From sequence to chromosome: the tip of the X chromosome of D.
 RT melanogaster.";
 RL Science 287:2220-2222(2000).
 CC -1- FUNCTION: NEUROGENIC PROTEIN ESSENTIAL FOR THE DEVELOPMENT AND
 CC MAINTENANCE OF EPITHELIAL STRUCTURE. REQUIRED IN THE GERMLINE FOR
 CC ESTABLISHING THE FOLLICULAR EPITHELIUM AND FOR DETERMINING THE
 CC DORSAL-VENTRAL POLARITY. COLLABORATES WITH NOTCH ON THE APICAL
 CC SURFACE OF FOLLICLE CELLS TO MEDIATE GERMLINE-FOLLICLE CELL
 CC ADHESION. BRN HAS A ROLE IN CHORION FORMATION.
 DR EMBL: U14449; AA05211.1; -.
 DR EMBL: AL033125; CAA21833.1; -.
 DR Flybase; FBgn0000221; brn.
 DR InterPro; IPR002659; Glyco_trans_31.
 DR Pfam; PF01762; Galactosyl_T; 1.
 DR Developmental protein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 325
 FT CONFLICT 142 142 D -> B (IN REF. 2).
 FT CONFLICT 163 163 E -> D (IN REF. 2).
 SQ SEQUENCE 325 AA; 37619 MW; BDA96F2DE04D7ECB CRC64;
 Query Match 58.8%; Score 47; DB 5; Length 325;
 Best Local Similarity 58.3%; Pred. No. 6.1;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DKMEAYITPGAF 12
 Db 229 DRWPVVTGAF 240
 ID 097N98 PRELIMINARY; PRT; 383 AA.
 AC 097N98;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Alcohol dehydrogenase, iron-containing.
 GN SP2157.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ITGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RX Letellin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gilm M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickenson T., Hickey B.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 DR EMBL; AE007504; AAK76211.1; -.
 DR TIGR; SP2157; -.
 DR InterPro; IPR001670; Fe-ADH.
 DR Pfam; PF00465; Fe-ADH; 1.
 DR PROSITE; PS00913; ADH_IRON_1; 1.
 DR PROSITE; PS00060; ADH_IRON_2; 1.
 KM Complete proteome.
 SQ SEQUENCE 383 AA; 41224 MW; E65A0C14E951B7BF CRC64;

Query Match 57.5%; Score 46; DB 16; Length 383;
 Best Local Similarity 72.7%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 4 EAYITPGAFDV 14
 Db 199 EALVTPGAYDV 209
 ID 08DN25 PRELIMINARY; PRT; 383 AA.
 AC 08DN25;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Probable alcohol dehydrogenase (EC 1.1.1.1).
 GN ADH2 OR SPRI963.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RX Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA Leblanc D.J., Lee L.N., Lefkowitz E.O., Lu J., Matsushima P.,
 RA McLaren S.M., McHenry M., McKeaster K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Roestek P.R. Jr., Skatrud P.L.,
 RA Glas J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL; AE008560; AAL00765.1; -.
 KM Oxidoreductase; Complete proteome.
 SQ SEQUENCE 383 AA; 41154 MW; 8E181FCD5A1684B CRC64;
 Query Match 57.5%; Score 46; DB 16; Length 383;
 Best Local Similarity 72.7%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 4 EAYITPGAFDV 14
 Db 199 EALVTPGAYDV 209
 ID 08DN25 PRELIMINARY; PRT; 765 AA.
 AC P74599;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein slr1567.
 GN SLR1567.
 OS Synchocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 OX NCBI_TaxID=1148;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RX Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Kamekura M., Hirose M., Sugita M., Saito S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).

DR EMBL: D90916; BAA18707.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 765 AA; 85407 MW; 97F8B3BD4800902 CRC64;

Query Match 56.2%; Score 45; DB 16; Length 765;
Best Local Similarity 63.6%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKMEAYITPGA 11
| | | | |
Db 160 DGMQIFITPGA 170

RESULT 8

OSDV3 PRELIMINARY; PRT; 169 AA.

AC OSDV3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE DUTPase.
OS Bacteriophage phi-11.
OC Viruses.
OX NCBI_TaxID=12360;
RN [1]
RP SEQUENCE FROM N.A.
RA Iandolo J.J., Worrell V., Roe B., Qian Y., Dorman A., Tian R., Lin S.,
RT Uta H.;
RT "Comparative analysis of the genomes of the temperate bacteriophages
RT #11, #12 and #13 of *Staphylococcus aureus* 8325."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF424781; ALR82253.1; -
DR InterPro: IPR001428; DeoxyUTPase.
DR Pfam: PF00692; DUTPase; 1.
DR ProDom: PD000946; DeoxyUTPase; 1.
DR TIGRFAMs: TIGR00576; duc; 1.
SQ SEQUENCE 169 AA; 18369 MW; 294A8624F333DAF CRC64;

Query Match 55.0%; Score 44; DB 9; Length 169;
Best Local Similarity 77.8%; Pred. No. 9.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 YITPGAFDV 14
| | | | |
Db 102 YITPGVFDI 110

RESULT 9

OSDV3 PRELIMINARY; PRT; 175 AA.

AC OSDV3;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DUTPase.
OS *Staphylococcus aureus* bacteriophage PVL.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=71366;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Muramoto K., Kamio Y.;
RT "Gene of *LukE*-*PV*-like component of Panton-Valentine leukocidin in
RT *Staphylococcus aureus* P83 is linked with *lukM*."
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98332719; PubMed=9666077;
RA Kaneko J., Kimura T., Narita S., Tomita T., Kamio Y.;
RT "Complete nucleotide sequence and molecular characterization of the
RT temperate *staphylococcal* bacteriophage phi PVL carrying Panton-
RT Valentine leukocidin genes."
RL Gene 215:57-67(1998).

DR EMBL: AB009866; BAA31926.1; -

DR HSSP: P16088; 1DUY.

DR InterPro: IPR001428; DeoxyUTPase.

DR Pfam: PF00692; DUTPase; 1.

DR ProDom: PD000946; DeoxyUTPase; 1.

DR TIGRFAMs: TIGR00576; duc; 1.

SQ SEQUENCE 175 AA; 19081 MW; B93B547104EC9BA0 CRC64;

Query Match 55.0%; Score 44; DB 9; Length 175;
Best Local Similarity 77.8%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 YITPGAFDV 14
| | | | |
Db 108 YITPGVFDI 116

RESULT 10

OSVCTS PRELIMINARY; PRT; 354 AA.

AC OSVCTS;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE CG4704 protein.
GN CG4704.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brodeur P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jajalji M., Kalush F., Kapen G.H., Ke Z., Kension J.A., Letchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrelti A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacble J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).

DR EMBL, AE003740; AAF56070.1; -
 DR FlyBase; FBgn0039029; CG4704.
 DR InterPro; IPR002048; EG-hand.
 DR Pfam; PF00036; ehand; 2.
 DR PROSITE; PS00018; EF_HAND; 2.
 KM Calcium; Calcium-binding.
 SQ SEQUENCE 354 AA; 41790 MW; 9B00F103BEC66B1 CRC64;

Query Match 55.0%; Score 44; DB 5; Length 354;
 Best Local Similarity 63.6%; Pred. No. 22;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DKWEAYITPGA 12
 Db 31 KWEIYMTPNDF 41

RESULT 11
 ID 070012 PRELIMINARY; PRT; 669 AA.

AC 070012;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Pepl.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinomycetales; Actinomycetaceae;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 RN NCBI_TaxID=1902;
 RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
 RX MEDLINE=20279210; PubMed=10821190;
 RA Schneider D., Bruton C.J., Chater K.F.;
 RT "Duplicate gene clusters suggest an interplay of glycogen and
 trehalose metabolism during sequential stages of aerial mycelium
 development in Streptomyces coeli color A3(2).";
 RL Mol. Gen. Genet. 263:543-553(2000).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=96154943; PubMed=8596463;
 RA Bruton C.J., Plaskitt K.A., Chater K.F.;
 RT "Tissue-specific glycogen branching isoenzymes in a multicellular
 prokaryote, Streptomyces coelicolor A3(2).";
 RL Mol. Microbiol. 18:89-99(1995).
 RN EMBL; A001206; CA04606.1; -

DR InterPro; IPR001589; Actbind_actnin.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR Pfam; PF00128; alpha-amy1ase; 1.
 DR PROSITE; PS00019; ACTNIN_1; 1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
 SQ SEQUENCE 669 AA; 74113 MW; 1563E0134CD16268 CRC64;

Query Match 55.0%; Score 44; DB 2; Length 669;
 Best Local Similarity 63.6%; Pred. No. 45;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEAYITPGA 11
 Db 77 DRWGATVTPGA 87

RESULT 12
 ID 09KY04 PRELIMINARY; PRT; 669 AA.

AC 09KY04;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative alpha-amy1ase.
 GN PEP1B OR SC07335 OR SC4G10.14C.

OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 RN NCBI_TaxID=1902;
 RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12009953;

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleiser T., Latke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 RN EMBL; AL939131; CAB92881.1; -

DR InterPro; IPR001589; Actbind_actnin.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR InterPro; IPR000531; TonB_boxC.
 DR Pfam; PF00128; alpha-amy1ase; 1.
 DR PROSITE; PS00019; ACTNIN_1; 1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
 KM Complete proteome.
 SQ SEQUENCE 669 AA; 74143 MW; 5463E0134CD66518 CRC64;

Query Match 55.0%; Score 44; DB 16; Length 669;
 Best Local Similarity 63.6%; Pred. No. 45;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEAYITPGA 11
 Db 77 DRWGATVTPGA 87

RESULT 13
 ID 043395 PRELIMINARY; PRT; 682 AA.

AC 043395;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE U4/U6-associated RNA splicing factor.
 GN PRP3.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=97472464; PubMed=9328476;
RA Wang A., Forman-Kay J., Luo Y., Luo M., Chow Y.-H., Plumb J.,
RA Priesen J.D., Tsui L.-C., Heng H.H.O., Woolford J.L. Jr., Hu J.;
RT Identification and characterization of human genes encoding Hprp3p
RL and Hprp4p, interacting components of the spliceosome.;
RL Hum. Mol. Genet. 6:2117-2126(1997).
DR EMBL; AF001947; AAC09069.1; -.
DR InterPro; IPR002483; PMI.
DR Pfam; PF01480; PMI.1.
DR SMART; SM00311; PMI.1.
SQ SEQUENCE 682 AA; 77403 MW; 544660BF72A39BA3 CRC64;

Query Match 55.0%; Score 44; DB 4; Length 682;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
DB 390 WDSYIIPNGFDL 401

RESULT 14
ID 043446 PRELIMINARY; PRT; 683 AA.
AC 043446;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE U4/U6 small nuclear ribonucleoprotein HPRP3 (U4/U6-associated RNA
DE splicing factor).
GN HPRP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98067393; PubMed=9404889;
RA Horowitz D.S., Kobayashi R., Krainer A.R.;
RT A new cyclophilin and the human homologues of yeast Prp3 and Prp4
RT form a complex associated with U4/U6 snRNPs.;
RL RNA 3:1374-1387(1997).
[2]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Skin, and Eye;
RA Strauberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016370; AAC51926.1; -.
DR EMBL; BC001954; AAH01954.1; -.
DR EMBL; BC000184; AAH00184.1; -.
DR InterPro; IPR002483; PMI.
DR Pfam; PF01480; PMI.1.
DR SMART; SM00311; PMI.1.
KW Nucleoprotein; Ribonucleoprotein.
SQ SEQUENCE 683 AA; 77528 MW; AAA6AA4C99110284 CRC64;

Query Match 55.0%; Score 44; DB 4; Length 683;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
DB 390 WDSYIIPNGFDL 401

RESULT 15
ID 09D6C6 PRELIMINARY; PRT; 683 AA.
AC 09D6C6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 3632413F13RIK protein.

```

```

GN 3632413F13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matano Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl R., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Oikido T., Furuno M., Aono H., Balarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK014398; BAB29324.1; -.
DR MGD; MGI:1918017; 3632413F13RIK.
DR InterPro; IPR000985; Lectin_Lega.
DR InterPro; IPR002483; PMI.
DR Pfam; PF01480; PMI.1.
DR SMART; SM00311; PMI.1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA.1.
SQ SEQUENCE 683 AA; 77430 MW; B1FC8E27DCE3DEFA CRC64;

Query Match 55.0%; Score 44; DB 11; Length 683;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
DB 390 WDSYIIPNGFDL 401

Search completed: January 12, 2004, 07:02:03
Job time : 33.4375 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 06:47:09 ; Search time 33.9062 Seconds
(without alignments)
65.539 Million cell updates/sec

Title: US-09-829-495-63

Perfect score: 80
Sequence: 1 DKMBAYITPGAFDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_19Jun03:*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	14	22	AA61295
2	80	100.0	14	24	ABU11261
3	52.5	65.6	267	23	AAU75503
4	47	58.8	193	23	AAE15968
5	47	58.8	325	22	ABE61108
6	46	57.5	383	24	ABE81494
7	46	57.5	383	24	ABU02685
8	45	56.2	831	22	ABG08487
9	44	55.0	169	18	AAW28034

10	44	55.0	354	22	ABB60901
11	44	55.0	369	24	ABU71061
12	44	55.0	867	22	AAE98887
13	43	53.8	207	21	AAE34190
14	43	53.8	217	21	AAE34189
15	43	53.8	223	21	AAE34188
16	43	53.8	424	22	ABE29934
17	43	53.8	906	22	ABE16321
18	43	53.8	1090	22	ABG17694
19	43	53.8	1090	22	ABG18331
20	42	52.5	71	22	AAO13076
21	42	52.5	243	23	ABE27382
22	41	51.2	172	22	AAU03641
23	41	51.2	205	22	AAU30764
24	41	51.2	245	23	ABE29803
25	41	51.2	249	23	ABE27381
26	41	51.2	312	21	AAE16153
27	41	51.2	363	18	AAW10320
28	41	51.2	378	16	AAE70205
29	41	51.2	378	17	AAW00621
30	41	51.2	378	23	ABE09592
31	41	51.2	378	23	ABE09593
32	41	51.2	406	18	AAE26604
33	41	51.2	411	16	AAE70206
34	41	51.2	420	16	AAE71326
35	41	51.2	1846	22	ABE63163
36	40	50.0	43	22	AAE90168
37	40	50.0	93	22	ABG04397
38	40	50.0	116	15	AAE52061
39	40	50.0	117	15	AAE52065
40	40	50.0	117	15	AAE52067
41	40	50.0	220	22	AAE04837
42	40	50.0	220	23	ABE69836
43	40	50.0	220	23	ABE32808
44	40	50.0	220	23	AAE14361
45	40	50.0	221	21	AAE18667

ALIGNMENTS

RESULT 1
AA61295
ID AAB61295 strand: Peptide; 14 AA.
XX
AC AAB61295;
XX
DT 04-APR-2001 (first entry)
XX
DE Anti-TANGO 268 scFv CDR, SEQ ID NO: 63.
XX
KW Human; antibody; scFv; CDR; complementarity determining region;
KW TANGO 268; cardiant; cerebroprotective; cytosolic; anticoagulant;
KW thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;
KW platelet membrane glycoprotein receptor; bleeding disorder;
KW blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;
KW ischemia; cardiovascular disease; immunological disease; liver disorder;
KW cancer.
XX
XX Homo sapiens.
XX
XX PN WO200100810-A1.
XX
XX PD 04-JAN-2001.
XX
XX PF 30-JUN-2000; 2000WO-US18152.
XX
XX PR 30-JUN-1999; 99US-0345468.
XX PR 06-DEC-1998; 98US-0454824.
XX PR 14-FEB-2000; 2000US-0503387.
XX
XX (MILL-) MILLENNIUM PHARM INC.

PI Busfield SJ, Villedal J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI Qian DM, Kingsbury G;
DR MPI; 2001-080877/09.

XX New genes encoding human platelet-expressed collagen receptor,
PT glycoprotein VI, and its modulators, useful for preventing, treating
PT and diagnosing hemorrhagic disorders, thrombotic diseases and
PT immunological disorders -

XX
PS Claim 31; Page 102; 227pp; English.

XX The present sequence is given in a specification relating to an isolated
CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
CC and polypeptides and their modulators, e.g. antisense nucleic acids,
CC ribozymes and antibodies, are useful for preventing, treating and
CC diagnosing disorders associated with aberrant expression or activity of
CC GPVI. These disorders include bleeding disorders
CC (e.g. thrombocytopenia), blood vessel injury, thrombotic disorders
CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC infarction), immunological diseases (e.g. platelet disorder) and
CC embryonic liver disorders. Preferably they are used to prevent acute
CC cardiac ischaemia following angioplasty and metastatic cancers,
CC especially of the colon and liver.

XX Sequence 14 AA;

Query Match 100.0%; Score 80; DB 22; Length 14;

Best Local Similarity 100.0%; Pred. No. 4.9e-07; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKWEAYITPGAPDV 14

Db 1 DKWEAYITPGAPDV 14

ABU11261 standard; Peptide; 14 AA.

AC ABU11261;

DT 06-FEB-2003 (first entry)

DE Human TANGO 268 VHCDR3 Peptide #3.

XX Human; mouse; variable heavy; VH; antigen: cancer;

KM complementarity determining region; TANGO 268; glycoprotein VI; GPVI;

KM TANGO 268; extracellular matrix; collagen; platelet release;

KM proliferation; migration; embryogenesis; inflammation; thrombosis;

KM degranulation; thrombocytopenia; antibody; thrombotic disorder;

KM cerebral vascular disease; stroke; ischaemia; venous thromboembolism;

KM leg swelling; pain; ulceration; pulmonary embolism; coronary disease;

KM cardiovascular disease; angina pectoris; myocardial infarction;

KM coronary restenosis; atherosclerosis; immunological disorder;

KM developmental disorder; embryonic disorder; liver disorder;

KM cerebral vascular disease; venous thromboembolism disease.

XX Homo sapiens.

OS WO200280968-A1.

XX 17-OCT-2002.

XX 09-APR-2002; 2002WO-US11122.

XX 09-APR-2001; 2001US-0829495.

XX (MILL-) MILLENNIUM PHARM INC.

PI Busfield SJ, Villedal J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI Qian DM, Kingsbury G;
DR MPI; 2003-058477/05.

XX Novel substantially purified antibody immunospecifically binding to
PT TANGO 268 antigen, useful for treating bleeding disorders such as
PT thrombocytopenia, stroke, ischaemia, pulmonary embolism, atherosclerosis
PT -

XX
PS Claim 8; Page 111; 236pp; English.

XX This invention relates to a novel purified antibody comprising a
CC variable heavy (VH) complementarity determining region (CDR)1, VH CDR2
CC or VH CDR3; or variable light (VL) CDR1, VL CDR2 or VL CDR3, and
CC immunospecifically binding to a TANGO 268 (also referred as glycoprotein
CC VI (GPVI)) antigen. The antibodies of the invention act to decrease or
CC block TANGO 268 binding to extracellular matrix components, or as a
CC collagen or platelet release and aggregation blocker. The antibodies of
CC the invention are useful for modulating proliferation, migration,
CC morphology, differentiation and/or function of megakaryocytes and
CC platelets, including during development e.g. embryogenesis, modulating
CC leukocyte-platelet and platelet-endothelium interactions in
CC inflammation and/or thrombosis, and modulating platelet aggregation and
CC degranulation. They are also useful for modulating disorders associated
CC with abnormal or aberrant megakaryocyte and/or platelet proliferation,
CC migration, morphology, differentiation and/or function, e.g. bleeding
CC disorders such as thrombocytopenia. Other diseases which may be
CC modulated by these antibodies are thrombotic disorders, cerebral
CC vascular diseases (e.g. stroke and ischaemia), venous thromboembolism
CC diseases (e.g. diseases involving leg swelling, pain and ulceration,
CC pulmonary embolism, etc); coronary diseases (e.g. cardiovascular
CC diseases including angina pectoris, myocardial infarction, coronary
CC restenosis, atherosclerosis, etc); immunological disorders,
CC developmental disorders, embryonic disorders, liver disorders, cerebral
CC vascular diseases, venous thromboembolism disease, coronary diseases,
CC and metastatic cancers. The antibodies of the invention only causes a
CC transient decrease in platelet counts, platelet aggregation, and/or
CC platelet activation and so have some advantages over prior art
CC methods. The present sequence represents a peptide sequence used to
CC generate the antibodies of the invention.

XX Sequence 14 AA;

Query Match 100.0%; Score 80; DB 24; Length 14;

Best Local Similarity 100.0%; Pred. No. 4.9e-07; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKWEAYITPGAPDV 14

Db 1 DKWEAYITPGAPDV 14

AAU75503 standard; Protein; 267 AA.

AC AAU75503;

DT 23-APR-2002 (first entry)

DE Human s-acyl fatty acid synthase thioesterase-like enzyme #2.

XX S-acyl fatty acid synthase thioesterase-like; SFTS-like; enzyme;

KM cardiac; anorectic; vasotrophic; extracellular matrix degradation;

KM cardiovascular disease; hyperlipidaemia; obesity; anorexia; cachexia;

KM wasting disorder; appetite suppression; appetite enhancement; bulimia;

KM diabetes; congestive heart failure; myocardial infarction; human;

KM ischemic disease; atrial arrhythmia; ventricular arrhythmia;

XX hypertensive vascular disease; peripheral vascular disease.

XX Homo sapiens.

PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers BW;
XX
DR WPI, 2001-656860/75.
DR N-PSDB; ABL05211.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 10116; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB16175), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;
XX
Query Match 58.8%; Score 47; DB 22; Length 325;
Best Local Similarity 58.3%; Pred. No. 7.6;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 DKWEAVITPGAF 12
|:|:|:|:|:
Db 229 DRMPYVTAGF 240
XX
RESULT 6
ABP81494
ID ABP81494 standard; Protein; 383 AA.
XX
AC ABP81494;
XX
DT 04-MAR-2003 (first entry)
XX
DE Streptococcus pneumoniae polypeptide SEQ ID NO 411.
XX
KW Streptococcus pneumoniae; infection; otitis media; antibacterial;
KW diagnosis; gene therapy.
XX
OS Streptococcus pneumoniae.
XX
PN WO200283855-A2.
XX
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-US11524.
XX
PR 16-APR-2001; 2001US-283948P.
PR 18-APR-2001; 2001US-28443P.
XX
PA (AMCY) AMERICAN CYANAMID CO.
XX
PI Zagursky RJ, Masi AW, Green BA, Chakravarti DN, Russell DP;
PI Wooleys JL;
XX
DR WPI, 2003-093010/08.
DR N-PSDB; ABZ42342.
XX
PT New Streptococcus pneumoniae polynucleotides, useful for treating or
PT preventing *S. pneumoniae* infections, or non-systemic diseases, e.g.
PT otitis media, which are induced or exacerbated by *S. pneumoniae* -
XX
PS Claim 42; Page 680-682; 1091pp; English.

XX
CC The invention relates to isolated polynucleotides (AB272147-AB242522) of
CC *Streptococcus pneumoniae* genomic sequence, a fragment or degenerate
CC variant of the polynucleotide or a nucleic acid sequence 95% identical to
CC one of the polynucleotides. The *S. pneumoniae* polynucleotides and
CC encoded polypeptides (ABP81299-ABP81674) are useful for treating or
CC preventing *S. pneumoniae* infections or non-systemic diseases, e.g. otitis
CC media, which are induced or exacerbated by *S. pneumoniae*. These are also
CC useful for detecting *S. pneumoniae* in a biological sample or diagnosing
CC *S. pneumoniae* infection in a subject. The polynucleotides have
CC antibacterial activity and are useful in gene therapy.
XX
SQ Sequence 383 AA;
XX
Query Match 57.5%; Score 46; DB 24; Length 383;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 4 EAVITPGADV 14
|:|:|:|:|:
Db 199 EALVTPGADV 209
XX
RESULT 7
ABU02685
ID ABU02685 standard; Protein; 383 AA.
XX
AC ABU02685;
XX
DT 11-FEB-2003 (first entry)
XX
DE *S. pneumoniae* type 4 strain protein from coding region #2264.
XX
KW Bacterial meningitis; pneumonia; sepsis; otitis media;
KW ear infection; antiinflammatory; antibacterial; immunostimulant;
KW auditory; respiratory; gene therapy; vaccine.
XX
OS Streptococcus pneumoniae type 4 strain.
XX
PN WO200277021-A2.
XX
PD 03-OCT-2002.
XX
PF 27-MAR-2002; 2002WO-IB02163.
XX
PR 27-MAR-2001; 2001GB-0007658.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Masignani V, Tettelin H, Fraser C;
PI N-PSDB; ABX07976.
XX
DR WPI, 2003-040579/03.
DR N-PSDB; ABX07976.
XX
PT New proteins and nucleic acid molecules from *Streptococcus pneumoniae*,
PT useful as medicaments for treating or preventing a disease or infection
PT due to *Streptococcus bacteria*, such as pneumonia, sepsis, otitis media
PT or ear infection -
XX
PS Claim 1; SEQ ID No 4528; 56pp; English.
XX
CC The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC *Streptococcus pneumoniae* type 4 strain genomic sequence appearing as
CC AB556454. Also included are an antibody which binds one of the
CC proteins, treating a patient by administering the protein, DNA or
CC antibody (in a composition), a kit comprising first and second primers,
CC which are the nucleic acid cited above or fragments between nucleotides
CC 8-100 of a sequence not defined in the specification, for amplifying a
CC target sequence contained within a *Streptococcus nucleic acid* sequence,

CC where the first primer is substantially complementary to the target
CC sequence and the second primer is substantially complementary to the
CC complement of the target sequence, and where the parts of the primers
CC having substantial complementarity define the termini of the target
CC sequence to be amplified, assay comprising contacting a test compound
CC with the protein, and determining whether the test compound binds to the
CC protein and a Streptococcus pneumoniae bacterium, where one or more
CC genes encoding the proteins has been rendered inactive. The proteins,
CC nucleic acid molecules, antibody and compositions are useful as
CC medicaments for treating or preventing a disease or infection due to
CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
CC sepsis, otitis media or ear infection. They are also useful in developing
CC vaccines, diagnostics and antibiotics. The methods are useful for
CC identifying immunodominant proteins. The present sequence is one of
CC the 2469 proteins expressed by the identified coding regions from the
CC genomic sequence.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
CC
XX

SO Sequence 383 AA;

Query Match 57.5%; Score 46; DB 24; Length 383;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAYITPGAPDV 14
|||:||||:
DB 199 EALVTPGAYDV 209

RESULT 8
ABG08487
ID ABG08487 standard; Protein; 831 AA.

XX ABG08487;

XX 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8478.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW Food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

PF 30-MAR-2001, 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX MPI; 2001-639362/73.

DR N-PSDB; AAS72674.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits and to assess

CC biodiversity -
PS Claim 20; SEQ ID No 38846; 103pp; English.

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX

SO Sequence 831 AA;

Query Match 56.2%; Score 45; DB 22; Length 831;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KWEAVITPGAP 12
||||:||||:
DB 72 KWEASVPPGSF 82

RESULT 9
AAW28034
ID AAW28034 standard; Protein; 169 AA.

XX AAW28034;

XX 27-AUG-1998 (first entry)

DE Amino acid sequence of the deoxyuridine 5' triphosphatase homologue.

XX Staphylococcus aureus protein; ribozyme; antisense sequence; control;

KW Staphylococcal gene; regulatory element; bacterial gene expression;

KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;

XX toxic shock syndrome.

XX Staphylococcus aureus.

XX WO9730070-A1.

XX 21-AUG-1997.

PF 19-FEB-1997; 97WO-US02318.

XX 20-FEB-1996; 96US-0011888.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;

PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;

XX MPI; 1997-424969/39.

DR N-PSDB; AAT83988.

XX Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used

PT to isolate antimicrobial compounds, and in vaccines against S.

CC aureus infection
PS Claim 6; Page 421; 989pp; English.

XX The present sequence represents a Staphylococcus aureus protein,
CC which is believed to be a deoxyuridine 5' triphosphatase homologue.
CC The DNA sequence was isolated from a library of clones of S. aureus
CC WCUH 29 in Escherichia coli. The DNA sequence can be used in the

CC construction of ribozymes and antisense sequences to control the
CC expression of Staphylococcal genes. The DNA sequence is also useful as
CC a source of regulatory elements for the control of bacterial gene
CC expression. The present protein may be used to produce vaccines to
CC enable a host to produce specific antibodies with antibacterial action.
CC These vaccines and antibodies would protect a host against invasion by
CC *S. aureus*, and conditions relating to staphylococcal infection, e.g.
CC Staphylococcal food poisoning, scaled skin syndrome, and toxic shock
CC syndrome.

XX SQ Sequence 169 AA;

Query Match 55.0%; Score 44; DB 18; Length 169;

Best Local Similarity 77.8%; Pred. No. 12;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 YITPGADV 14
|||
Db 102 YITPGVFDI 110

RESULT 10

ABB60901 ABB60901 standard; Protein; 354 AA.

XX ABB60901;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 9495.

XX Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li FMD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL05004.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure; SEQ ID NO 9495; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511); expressed DNA
XX sequences (ABL01840-AB116175) and the encoded proteins
XX (ABB57731-ABB720722).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 354 AA;

Query Match 55.0%; Score 44; DB 22; Length 354;

Best Local Similarity 63.6%; Pred. No. 27;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAVITPGAF 12
|||
Db 31 KWEAVITPNDP 41

RESULT 11

ABU71061 ABU71061 standard; Protein; 369 AA.

XX ABU71061;

XX 10-JUN-2003 (first entry)

XX Human adipocyte Selected Interacting domain, SID, #692.

XX Human; prey; adipocyte; SID; selected interacting domain;

XX anorectic; antidiabetic; protein-protein interaction; diabetes;

XX Yeast 2-hybrid assay; metabolic disorder; obesity.

XX Homo sapiens.

XX WO200286122-A2.

XX 31-OCT-2002.

XX 14-MAR-2002; 2002WO-EP03768.

XX 14-MAR-2001; 2001US-275734P.

XX (HYBR-) HYBRIGENICS.

XX Legrain P, Daviet L;

XX WPI; 2003-103412/09.

XX N-PSDB; ACA57605.

XX New complex between two interacting proteins in adipocyte cells, useful
XX for identifying selected interacting domains that modulate protein
XX interactions, or for preventing or treating metabolic disorders such as
XX obesity or diabetes -

XX Claim 6; Page 346; 382pp; English.

XX The invention relates to a complex between two interacting proteins in
XX adipocyte cells, given in the specification. The proteins are identified
XX by selecting a bait protein from a known adipocyte marker and then
XX performing a yeast 2-hybrid selection to isolate prey proteins encoded by
XX members of an adipocyte cDNA library. The proteins are designated SID
XX (RTM) (selected interacting domain) proteins. Also included are a
XX polynucleotide encoding a polypeptide in the adipocyte cells, a
XX recombinant host cell expressing at least one of the interacting
XX polypeptides of the complex, selecting a modulating compound in adipocyte
XX cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
XX sequences given in the specification (including its fragment or variant),
XX a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
XX given in the specification (including its fragment or variant), a vector
XX comprising the SID (RTM) polynucleotide, a recombinant host cell
XX comprising the vector, a protein chip comprising the polypeptides and
XX a record comprising all or part of the data, listed in the specification.
XX The complex, polypeptides, polynucleotides and compounds are
XX useful for preventing or treating metabolic disorders such as obesity
XX or diabetes. The polynucleotides are useful as probes or primers. The
XX complex is particularly useful for identifying selected interacting
XX domains (SID (RTM)) for screening drugs that modulate the protein
XX interaction, thus exhibiting the therapeutic effect. The present
XX sequence represents a SID (prey) protein of the invention.

XX SQ Sequence 369 AA;

Query Match 55.0%; Score 44; DB 24; Length 369;

Best Local Similarity: 50.0%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14
|::|||::|
237 WDSYITPNGFDL 248

RESULT 12
AAG98887
ID AAG98887 standard; Protein; 867 AA.

AC AAG98887;

DT 26-SEP-2001 (first entry)

DE E. coli growth and proliferation related protein sequence SEQ ID NO:357.

KW Escherichia coli; growth; proliferation; microbial; antimicrobial;
bacterial infection; microorganism.

OS Escherichia coli.

PN WO200134810-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30950.

PR 09-NOV-1999; 99US-0164415.

PA (ELITR-) ELITRA PHARM INC.

PI Forsyth RA, Ohlsen K, Zyskind J;

XX WPI, 2001-335933/35.

DR N-PSDB; AAH84558.

PT Novel nucleic acids that inhibit Escherichia coli proliferation, useful
for screening for homologous genes and for designing expression vectors

PS Claim 19; Page 432-434; 522pp; English.

AAH84373 to AAH84499 represent Escherichia coli growth and proliferation
related DNA sequences (I). AAH84500 to AAH84670 encode the E. coli
growth and proliferation related proteins given in AAG99078 and AAG98830
to AAG98999. (I) can be used as potential targets for the generation of
new antimicrobial agents, and for identification of compounds which
interact with the gene products of (I). In addition the expression of
(I) and the purification of the proteins, the purified proteins can be
used to generate reagents and screen small molecule libraries or other
candidate compound libraries for compounds that can be further developed
to yield novel antimicrobial compounds. In addition, nucleic acid probes
complementary to (I) that are specific for particular species of
microorganisms can be used to identify particular microorganism species
in clinical specimens, therefore, providing a rapid and dependable
method by which to identify the causative agents of a bacterial
infection. Also, antibodies generated against proteins translated from
mRNA transcribed from proliferation-related sequences can also be used
to screen for specific microorganisms that produce such proteins in a
species-specific manner. AAH84371 and AAH84670 represent sequencing
primers used in the isolation of E. coli growth and proliferation
related sequence, which are used in an example from the present
invention.

XX Sequence 867 AA;

Query Match 55.0%; Score 44; DB 22; Length 867;
Best Local Similarity 41.7%; Pred. No. 73;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 3 WEAYITPGAFDV 14

DB 309 YOSVSPGAFEI 320
:::||||:::

RESULT 13
AAG34190
ID AAG34190 standard; Protein; 207 AA.

AC AAG34190;

DT 18-OCT-2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 41561.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence; corn.

OS Zea mays subsp. mays.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135623.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 24-JUN-1999; 99US-0140656.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142350.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0135542.
PR 14-JUL-1999; 99US-0135624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161044.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.8%; Score 43; DB 21; Length 207;
Best Local Similarity 42.9%; Pred. No. 23;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAPDV 14
:|||||:|:|:
Db 83 EWEWAILTPEOFNI 96

RESULT 14
AAG34189
ID AAG34189 standard; Protein; 217 AA.

```
XX AC AAG34189;
XX 18-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 41560.
XX
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX OS Zea mays subsp. mays.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123348.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130549.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134421.
XX PR 18-MAY-1999; 99US-0134470.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 17-JUN-1999; 99US-0139453.
XX PR 17-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
```

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 22-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157855.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.8%; Score 43; DB 21; Length 217;
Best Local Similarity 42.9%; Pred. No. 24;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 DKWEAYITPGAFDV 14
::|||::|||::
DB 93 EWEAILTPQEFNI 106

RESULT 15
AAG34188
ID AAG34188 standard; Protein; 223 AA.

XX AAG34188;
XX AC
XX DT 18-OCT-2000 (first entry)
XX DE
XX Zee mays protein fragment SEQ ID NO: 41559.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
OS Zea mays subsp. mays.
XX EPI033405-A2.
XX PD
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 18-MAY-1999; 99US-0134841.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136392.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139452.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147039.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.8%; Score 43; DB 21; Length 223;
Best Local Similarity 42.9%; Pred. No. 25;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 DKWEAYITPGADV 14
DB 99 EWEWAILTPEQFNI 112

Search completed: January 12, 2004, 06:58:54
Job time : 35.9062 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 07:02:25 ; Search time 28 Seconds

(without alignments)
100.755 Million cell updates/sec

Title: US-09-829-495-63

Sequence: 1 DKWEAYITPGARDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications_AA.*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	14	9 US-09-832-312-63	Sequence 63, Appl
2	80	100.0	14	12 US-09-829-495-63	Sequence 63, Appl
3	47	58.8	333	10 US-09-972-912-3	Sequence 3, Appl
4	47	58.8	324	16 US-10-344-440-3	Sequence 3, Appl
5	47	58.8	325	9 US-09-804-357-10	Sequence 10, Appl
6	47	58.8	325	9 US-09-739-451-5	Sequence 5, Appl
7	47	58.8	325	10 US-09-804-006-10	Sequence 10, Appl
8	47	58.8	325	11 US-09-284-320-90	Sequence 90, Appl
9	44	55.0	867	12 US-10-287-274-357	Sequence 357, Appl
10	43	53.8	333	12 US-10-369-493-18135	Sequence 18135, A
11	41	51.2	172	12 US-10-091-007-82	Sequence 82, Appl
12	40	50.0	180	10 US-09-847-519A-17	Sequence 17, Appl
13	40	50.0	220	10 US-09-847-519A-2	Sequence 2, Appl
14	40	50.0	868	12 US-10-369-493-3985	Sequence 3985, Ap
15	39	48.8	141	9 US-09-847-539A-5	Sequence 5, Appl

16	39	48.8	155	9 US-09-847-539A-9	Sequence 9, Appl
17	39	48.8	159	9 US-09-847-539A-6	Sequence 6, Appl
18	39	48.8	167	9 US-09-847-539A-11	Sequence 11, Appl
19	39	48.8	184	9 US-09-847-539A-4	Sequence 4, Appl
20	39	48.8	217	9 US-09-847-539A-7	Sequence 7, Appl
21	39	48.8	264	16 US-10-080-170-396	Sequence 396, Appl
22	39	48.8	271	9 US-09-847-539A-10	Sequence 10, Appl
23	39	48.8	397	12 US-10-369-493-16747	Sequence 16747, A
24	38	47.5	117	12 US-10-255-478-45	Sequence 45, Appl
25	38	47.5	133	12 US-10-255-478-48	Sequence 48, Appl
26	38	47.5	134	12 US-10-255-478-49	Sequence 49, Appl
27	38	47.5	150	12 US-09-226-157-2	Sequence 2, Appl
28	38	47.5	264	16 US-10-080-170-289	Sequence 289, Appl
29	38	47.5	271	12 US-09-226-157-4	Sequence 4, Appl
30	38	47.5	279	15 US-10-156-761-10577	Sequence 10577, A
31	38	47.5	322	15 US-10-156-761-14966	Sequence 14966, A
32	38	47.5	7257	14 US-10-014-717-5	Sequence 5, Appl
33	37.5	46.9	69	15 US-10-156-761-11318	Sequence 11318, A
34	37	46.2	63	12 US-10-076-747-108	Sequence 108, Appl
35	37	46.2	274	12 US-10-369-493-3041	Sequence 3041, Ap
36	37	46.2	345	12 US-10-032-585-7871	Sequence 7871, Ap
37	37	46.2	391	12 US-10-080-334-227	Sequence 227, Appl
38	37	46.2	447	15 US-10-156-761-14785	Sequence 14785, A
39	37	46.2	494	15 US-09-738-626-5829	Sequence 6829, Ap
40	37	46.2	510	9 US-09-925-301-1131	Sequence 1131, Ap
41	37	46.2	2374	11 US-09-383-894-2	Sequence 2, Appl
42	37	46.2	2425	11 US-09-383-894-4	Sequence 4, Appl
43	36	45.0	190	10 US-09-764-864-1071	Sequence 1071, Ap
44	36	45.0	209	9 US-09-815-242-5452	Sequence 5452, Ap
45	36	45.0	209	9 US-09-815-242-12382	Sequence 12382, A

ALIGNMENTS

RESULT 1
US-09-832-312-63
; Sequence 63, Application US/09832312
; Patent No. US20010049829A1
GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-63
Query Match 100.0%; Score 80; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DKWEAYITPGARDV 14
|||||
Db 1 DKWEAYITPGARDV 14
RESULT 2
US-09-829-495-63
; Sequence 63, Application US/09829495

Publication No. US20040001826A1
GENERAL INFORMATION:
APPLICANT: Buefield SJ
APPLICANT: Valleval J
APPLICANT: Vandroet-Petrus M
APPLICANT: Vainchenker W
APPLICANT: G111 DS
APPLICANT: Qian MD
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-234
CURRENT APPLICATION NUMBER: US/09/829,495
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 09/610,118
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/503,387
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 09/454,824
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 09/345,468
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 63
LENGTH: 14
TYPE: PRT
ORGANISM: Homo sapiens
US-09-829-495-63

Query Match 100.0%; Score 80; DB 12; Length 14;
Best Local Similarity 100.0%; Pred. No. 1,5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAPDV 14
DB 1 DKWEAYITPGAPDV 14

RESULT 3

US-09-972-912-3

Sequence 3, Application US/09972912
Patent No. US20020110867A1
GENERAL INFORMATION:
APPLICANT: SOPPET, DANIEL R.
RUBEN, STEVEN M.
TITLE OF INVENTION: CARDIAC AND PANCREATIC PROTEIN AND GENE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: US
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,912
FILING DATE: 10-OCT-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/049,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: STEEFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488,0620001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 323 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020110867A1 Relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-972-912-3

Query Match 58.8%; Score 47; DB 10; Length 323;
Best Local Similarity 58.3%; Pred. No. 8.9;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
DB 228 DRMPYVTAGAF 239

RESULT 4

US-10-344-440-3
Sequence 3, Application US/10344440
Publication No. US20030131378A1
GENERAL INFORMATION:

APPLICANT: Aroian, Raffi
TITLE OF INVENTION: METHODS FOR BLOCKING RESISTANCE TO Bc TOXINS IN INSECTS AND NEMATODES
FILE REFERENCE: 6627-PA1023
CURRENT APPLICATION NUMBER: US/10/344,440
CURRENT FILING DATE: 2003-02-10
PRIOR APPLICATION NUMBER: 60/224,941
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: PCT/US01/41687
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 324
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-10-344-440-3

Query Match 58.8%; Score 47; DB 16; Length 324;
Best Local Similarity 58.3%; Pred. No. 8.9;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
DB 229 DRMPYVTAGAF 240

RESULT 5

US-09-804-357-10

Sequence 10, Application US/09804357
Patent No. US20010024808A1
GENERAL INFORMATION:
APPLICANT: White, David
APPLICANT: Zhou, Jiaohong
APPLICANT: Tartagila, Louis A.
TITLE OF INVENTION: LEPTIN INDUCED GENES
FILE REFERENCE: 07334/109001
CURRENT APPLICATION NUMBER: US/09/804,357
CURRENT FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 09/195,896
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: US 60/108,379
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: US 09/150,857
PRIOR FILING DATE: 1998-09-10
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 325
TYPE: PRT
ORGANISM: Drosophila melanogaster

US-09-804-357-10

Query Match	58.8%	Score 47;	DB 9;	Length 325
Best Local Similarity	58.3%;	Pred. No. 9;		
Matches	7;	Conservative	2;	Mismatches 3; Indels

```
QY      1 DKWEAYITPGAF 12
          | : | | : | |
Db      229 DRWPPYVTAGAF 240
```

RESULT 6
US-09-739-451-5

Query Match	58.8%;	Score 47;	DB 9;	Length 325;
Best Local Similarity	58.3%;	Pred. No. 9;		
Matches	7;	Conservative	2;	Mismatches 3;
				Indels 0;
				Gaps 0

```
QY      1 DKWEAYITPGAF 12
        | : | | |
Db      229 DRWPPYVTAGAF 240
```

```

RESULT 7
US-09-804-006-10
: Sequence 10, Application US/09804006
: Patent No. US20020119517A1
: GENERAL INFORMATION:
: APPLICANT: White, David
: APPLICANT: Zhou, Jianghong
: APPLICANT: Tartaglia, Louis A.
: TITLE OF INVENTION: LEFTIN INDUCED GENES
: FILE REFERENCE: 0734/4126001
: CURRENT APPLICATION NUMBER: US/09/804, 006
: CURRENT FILING DATE: 2001-03-12
: PRIOR APPLICATION NUMBER: US 09/292,228
: PRIOR FILING DATE: 1999-04-15
: PRIOR APPLICATION NUMBER: US 60/108,379
: PRIOR FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: US 09/150,857
: PRIOR FILING DATE: 1998-09-10
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: FASTSEQ for Windows Version 3.0.
: SEQ ID NO 10
: LENGTH: 325
: TYPE: prt
: ORGANISM: Drosophila melanogaster
US-09-804-006-10

```

Query Match	58.8%;	Score 47;	DB 10;	Length 325;
Best Local Similarity	58.3%;	Pred. No. 9;		
Matches	7;	Conservative	2;	Mismatches 3; Indels 0; Gaps 0

```
QY      1 DKWEAYITPGAF 12
          | : | | | |
Db      229 DRWPPYVTAGAF 240
```

```

RESULT 8 .
US-09-284-320-90
; Sequence 90. Application US/09284320
; Publication No. US20030092175A1
; GENERAL INFORMATION:
; APPLICANT: Kato, Seishi et al.
; TITLE OF INVENTION: HUMAN PROTEINS HAVING TRANSMEMBRANE DOMAINS AND DNAs
; TITLE OF INVENTION: ENCODING THESE PROTEINS
; FILE REFERENCE: GIN-6705CPUS
; CURRENT APPLICATION NUMBER: US/09/284,320
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: JP 8-301429
; PRIOR FILING DATE: 1996-11-13
; PRIOR APPLICATION NUMBER: PCT/JP97/04056
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila sp.
US-09-284-320-90

```

Query Match	58.8%	Score 47;	DB 11;	Length 325;
Best Local Similarity	58.3%	Pred. No. 9;		
Matches	7;	Conservative	2;	Mismatches 3; Indels 0; Gaps 0

QY	1	DKWEAYITPGAF	12
		: :	
Db	229	DRWPPYVTAGAF	240

```

RESULT 9
US-10-287-274-357
; Sequence 357, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; FILE REFERENCE: ELITRA.008DVI
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
; LENGTH: 867
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-357

```

Query Match	55.0%;	Score 44;	DB 12;	Length 867;
Best Local Similarity	41.7%;	Pred. No. 73;		
Matches	5;	Conservative	7;	Mismatches 0;
			Indels	0;
			Gaps	0;

```
QY      3 WEAYITPGAEDV 14  
          ::::|::|::  
Db      309 YQSYVSPGAFFI 320
```

RESULT 10
US-10-369-493-18135

```
/ Sequence 18135, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 18135
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Thermoplasma acidophilum
/ US-10-369-493-18135

Query Match      53.8%; Score 43; DB 12; Length 333;
Best Local Similarity 38.5%; Pred. No. 42;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFD 13
DB 199 DSWAGYLSPGBYE 211

RESULT 11
US-10-091-007-82
/ Sequence 82, Application US/10091007
/ Publication No. US20030170782A1
/ GENERAL INFORMATION:
/ APPLICANT: Microbial Techniques limited
/ APPLICANT: Le Page, Richard W F Wells, Jeremy M
/ APPLICANT: Hanniffy, Sean B
/ TITLE OF INVENTION: Proteins
/ FILE REFERENCE: PWC/P21978WO
/ CURRENT APPLICATION NUMBER: US/10/091,007
/ CURRENT FILING DATE: 2002-03-06
/ PRIOR APPLICATION NUMBER: GB 9921125.2
/ PRIOR FILING DATE: 1999-09-07
/ NUMBER OF SEQ ID NOS: 276
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 82
/ LENGTH: 172
/ TYPE: PRT
/ ORGANISM: Streptococcus agalactiae
/ US-10-091-007-82

Query Match      51.2%; Score 41; DB 12; Length 172;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 YITPGAFD 13
DB 92 YITPGSFD 99

RESULT 12
US-09-847-519A-17
/ Sequence 17, Application US/09847519A
/ Patent No. US20020102693A1
/ GENERAL INFORMATION:
/ APPLICANT: Lucbe, Ralf M.
/ APPLICANT: Wei, Bo
/ TITLE OF INVENTION: DSP-14 DUAL-SPECIFICITY PHOSPHATASE
/ FILE REFERENCE: 200125.422
/ CURRENT APPLICATION NUMBER: US/09/847,519A
/ CURRENT FILING DATE: 2001-05-01
```

```
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 17
/ LENGTH: 180
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-847-519A-17

Query Match      50.0%; Score 40; DB 10; Length 180;
Best Local Similarity 63.6%; Pred. No. 71;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAYITPGAFDV 14
DB 5 EDYCTPGAFEL 15

RESULT 13
US-09-847-519A-2
/ Sequence 2, Application US/09847519A
/ Patent No. US20020102693A1
/ GENERAL INFORMATION:
/ APPLICANT: Lucbe, Ralf M.
/ APPLICANT: Wei, Bo
/ TITLE OF INVENTION: DSP-14 DUAL-SPECIFICITY PHOSPHATASE
/ FILE REFERENCE: 200125.422
/ CURRENT APPLICATION NUMBER: US/09/847,519A
/ CURRENT FILING DATE: 2001-05-01
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 220
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-847-519A-2

Query Match      50.0%; Score 40; DB 10; Length 220;
Best Local Similarity 63.6%; Pred. No. 87;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAYITPGAFDV 14
DB 31 EDYCTPGAFEL 41

RESULT 14
US-10-369-493-3985
/ Sequence 3985, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 3985
/ LENGTH: 868
/ TYPE: PRT
/ ORGANISM: Neurospora crassa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1) - (868)
/ OTHER INFORMATION: unsure at all Xaa locations
/ US-10-369-493-3985
```

Query Match 50.0%; Score 40; DB 12; Length 868;
 Best Local Similarity 63.6%; Pred. No. 3.3e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KWEAYITPGAF 12
 : ||||| :
 Db 246 RWFAYITIGSF 256

RESULT 15

US-09-847-539A-5
 ; Sequence 5, Application US/09847539A
 ; Patent No. US20020061306A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bjorck, Lars H
 ; APPLICANT: Rasmussen, Magnus
 ; TITLE OF INVENTION: STREPTOCOCCAL ALPHA ZM BINDING PROTEIN
 ; FILE REFERENCE: 100084.415US / N.75312B
 ; CURRENT APPLICATION NUMBER: US/09/847,539A
 ; CURRENT FILING DATE: 2001-05-01
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 141
 ; TYPE: PR1
 ; ORGANISM: Streptococcus pyogenes
 US-09-847-539A-5

Query Match 48.8%; Score 39; DB 9; Length 141;
 Best Local Similarity 57.1%; Pred. No. 82;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFDV 14
 : ||||| :
 Db 116 DAMEKATPIALDV 129

Search completed: January 12, 2004, 07:19:54
 Job time : 30 secs

THIS PAGE BLANK (USPTO)